Zeman, R.
09/284787
Seq. IDs 1-3 w/ Interf

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 11, 2002, 10:14:10 ; Search time 266.73 Seconds
(without alignments)
9.369 Million cell updates/sec

Title: US-09-284-787-1
Perfect score: 55
Sequence: 1 YPYDVPDYA 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 3148936 seqs, 277657034 residues

Total number of hits satisfying chosen parameters: 3148936

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending_Patents_AA_Main:*
1: /cgn2_6/ptodata/2/paa/PCTUS_COMB.pep:*
2: /cgn2_6/ptodata/2/paa/US06_COMB.pep:*
3: /cgn2_6/ptodata/2/paa/US07_COMB.pep:*
4: /cgn2_6/ptodata/2/paa/US080_COMB.pep:*
5: /cgn2_6/ptodata/2/paa/US081_COMB.pep:*
6: /cgn2_6/ptodata/2/paa/US082_COMB.pep:*
7: /cgn2_6/ptodata/2/paa/US083_COMB.pep:*
8: /cgn2_6/ptodata/2/paa/US084_COMB.pep:*
9: /cgn2_6/ptodata/2/paa/US085_COMB.pep:*
10: /cgn2_6/ptodata/2/paa/US086_COMB.pep:*
11: /cgn2_6/ptodata/2/paa/US087_COMB.pep:*
12: /cgn2_6/ptodata/2/paa/US088_COMB.pep:*
13: /cgn2_6/ptodata/2/paa/US089_COMB.pep:*
14: /cgn2_6/ptodata/2/paa/US090_COMB.pep:*
15: /cgn2_6/ptodata/2/paa/US091_COMB.pep:*
16: /cgn2_6/ptodata/2/paa/US092_COMB.pep:*
17: /cgn2_6/ptodata/2/paa/US093_COMB.pep:*
18: /cgn2_6/ptodata/2/paa/US094_COMB.pep:*
19: /cgn2_6/ptodata/2/paa/US095_COMB.pep:*
20: /cgn2_6/ptodata/2/paa/US096_COMB.pep:*
21: /cgn2_6/ptodata/2/paa/US097_COMB.pep:*
22: /cgn2_6/ptodata/2/paa/US098_COMB.pep:*
23: /cgn2_6/ptodata/2/paa/US099_COMB.pep:*
24: /cgn2_6/ptodata/2/paa/US60_COMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	55	100.0	9	1	PCT-US00-05097-17	Sequence 17, Appl
2	55	100.0	9	1	PCT-US00-06950-9	Sequence 9, Appli
3	55	100.0	9	1	PCT-US00-19198-20	Sequence 20, Appl
4	55	100.0	9	1	PCT-US00-20131A-28	Sequence 28, Appl
5	55	100.0	9	1	PCT-US00-20131A-43	Sequence 43, Appl
6	55	100.0	9	1	PCT-US00-22456-1	Sequence 1, Appli
7	55	100.0	9	1	PCT-US00-23158-10	Sequence 10, Appl
8	55	100.0	9	1	PCT-US00-30238-5	Sequence 5, Appli
9	55	100.0	9	1	PCT-US93-12643-6	Sequence 6, Appli
10	55	100.0	9	1	PCT-US95-01467-14	Sequence 14, Appl
11	55	100.0	9	1	PCT-US97-15500-12	Sequence 12, Appl
12	55	100.0	9	1	PCT-US97-20344-11	Sequence 11, Appl
13	55	100.0	9	1	PCT-US97-22198-1	Sequence 1, Appli
14	55	100.0	9	1	PCT-US98-06791-1	Sequence 1, Appli
15	55	100.0	9	1	PCT-US98-11189-28	Sequence 28, Appl
16	55	100.0	9	1	PCT-US98-11189-28	Sequence 28, Appl
17	55	100.0	9	1	PCT-US99-01188-1	Sequence 1, Appli
18	55	100.0	9	1	PCT-US99-03085-93	Sequence 93, Appl
19	55	100.0	9	1	PCT-US99-04548-25	Sequence 25, Appl
20	55	100.0	9	4	US-08-017-931A-18	Sequence 18, Appl
21	55	100.0	9	4	US-08-092-977-19	Sequence 19, Appl
22	55	100.0	9	4	US-08-093-499-19	Sequence 19, Appl
23	55	100.0	9	5	US-08-108-351-7	Sequence 7, Appli
24	55	100.0	9	5	US-08-179-748-19	Sequence 19, Appl
25	55	100.0	9	5	US-08-192-477-14	Sequence 14, Appl
26	55	100.0	9	5	US-08-196-043-19	Sequence 19, Appl
27	55	100.0	9	6	US-08-287-537-3	Sequence 3, Appli
28	55	100.0	9	6	US-08-292-595-19	Sequence 19, Appl
29	55	100.0	9	6	US-08-292-596-19	Sequence 19, Appl
30	55	100.0	9	7	US-08-360-694-3	Sequence 3, Appli
31	55	100.0	9	8	US-08-443-982A-7	Sequence 7, Appli
32	55	100.0	9	8	US-08-457-263-3	Sequence 3, Appli
33	55	100.0	9	8	US-08-465-772A-28	Sequence 28, Appl
34	55	100.0	9	8	US-08-469-469A-28	Sequence 28, Appl
35	55	100.0	9	8	US-08-484-878-10	Sequence 10, Appl
36	55	100.0	9	10	US-08-619-484A-7	Sequence 7, Appli
37	55	100.0	9	10	US-08-642-120-31	Sequence 31, Appl
38	55	100.0	9	10	US-08-685-625-4	Sequence 4, Appli
39	55	100.0	9	10	US-08-690-011-17	Sequence 17, Appl
40	55	100.0	9	10	US-08-690-011-49	Sequence 49, Appl
41	55	100.0	9	11	US-08-750-111-4	Sequence 4, Appli
42	55	100.0	9	11	US-08-756-416-6	Sequence 6, Appli
43	55	100.0	9	11	US-08-793-076-54	Sequence 54, Appl
44	55	100.0	9	12	US-08-807-014-36	Sequence 36, Appl
45	55	100.0	9	24	US-60-291-149-39	Sequence 39, Appl

ALIGNMENTS

```
RESULT 1
PCT-US00-05097-17
; Sequence 17, Application PC/TUS0005097
; GENERAL INFORMATION:
;  APPLICANT:  Washington University
;  TITLE OF INVENTION:  NOVEL TRANSDUCTION MOLECULES AND METHODS
;  TITLE OF INVENTION:  FOR USING SAME
;  FILE REFERENCE:  49054-PCT
;  CURRENT APPLICATION NUMBER: PCT/US00/05097
;  CURRENT FILING DATE:  1999-02-28
;  PRIOR APPLICATION NUMBER:  60/122,757
;  PRIOR FILING DATE:  1999-02-28
;  PRIOR APPLICATION NUMBER:  60/151,291
;  PRIOR FILING DATE:  1999-08-29
;  NUMBER OF SEQ ID NOS:  54
;  SOFTWARE:  FastSEQ for Windows Version 3.0
; SEQ ID NO 17
;   LENGTH: 9
;   TYPE: PRT
;   ORGANISM: Unknown
;   FEATURE:
;   OTHER INFORMATION: Synthetic sequence
PCT-US00-05097-17


  Query Match           100.0%;  Score 55;  DB 1;  Length 9;
  Best Local Similarity  100.0%;  Pred. No. 2.9e+06;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 YPYDVPDYA 9
           |||||||||
Db       1 YPYDVPDYA 9
```

```
RESULT   2
PCT-US00-06950-9
; Sequence 9, Application PC/TUS0006950
; GENERAL INFORMATION:
;  APPLICANT:  VALIGENE CORPORATION
;  TITLE OF INVENTION:  METHODS FOR DETECTION OF NUCLEIC ACID POLYMORPHISMS USING
;  TITLE OF INVENTION:  PEPTIDE-LABELED OLIGONUCLEOTIDES AND ANTIBODY ARRAYS
;  FILE REFERENCE:  9408-042-228
;  CURRENT APPLICATION NUMBER: PCT/US00/06950
;  CURRENT FILING DATE:  2000-03-16
;  NUMBER OF SEQ ID NOS:  11
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 9
;   LENGTH: 9
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Oryctolagus cuniculus
PCT-US00-06950-9


  Query Match           100.0%;  Score 55;  DB 1;  Length 9;
  Best Local Similarity  100.0%;  Pred. No. 2.9e+06;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 YPYDVPDYA 9
          |||||||||
Db      1 YPYDVPDYA 9


RESULT   3
PCT-US00-19198-20
; Sequence 20, Application PC/TUS0019198
; GENERAL INFORMATION:
;  APPLICANT:  Human Genome Sciences, Inc.
;  TITLE OF INVENTION:  Follistatin-3
;  FILE REFERENCE:  PF388PCT2
;  CURRENT APPLICATION NUMBER: PCT/US00/19198
;  CURRENT FILING DATE:  2000-07-14
;  PRIOR APPLICATION NUMBER:  60/144,088
;  PRIOR FILING DATE:  1999-07-16
;  NUMBER OF SEQ ID NOS:  22
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 20
;   LENGTH: 9
;   TYPE: PRT
;   ORGANISM: Primer_Bind
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: hemaglutanin
;   OTHER INFORMATION: tag
PCT-US00-19198-20


  Query Match           100.0%;  Score 55;  DB 1;  Length 9;
  Best Local Similarity  100.0%;  Pred. No. 2.9e+06;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 YPYDVPDYA 9
          |||||||||
Db      1 YPYDVPDYA 9


RESULT   4
PCT-US00-20131A-28
; Sequence 28, Application PC/TUS0020131A
; GENERAL INFORMATION:
;  APPLICANT:  James S. HUSTON et al.
;  TITLE OF INVENTION:  METHODS AND COMPOSITIONS FOR INHIBITING POLYPEPTIDE
;  TITLE OF INVENTION:  ACCUMULATION ASSOCIATED WITH NEUROLOGICAL DISORDERS
;  FILE REFERENCE:  INR-004PC
;  CURRENT APPLICATION NUMBER: PCT/US00/20131A
;  CURRENT FILING DATE:  2000-07-24
;  PRIOR APPLICATION NUMBER:  60/146,047
;  PRIOR FILING DATE:  1999-07-27
;  NUMBER OF SEQ ID NOS:  45
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 28
;   LENGTH: 9
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: Synthetic
;   OTHER INFORMATION: construct
PCT-US00-20131A-28


  Query Match           100.0%;  Score 55;  DB 1;  Length 9;
  Best Local Similarity  100.0%;  Pred. No. 2.9e+06;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 YPYDVPDYA 9
          |||||||||
Db      1 YPYDVPDYA 9


RESULT   5
PCT-US00-20131A-43
; Sequence 43, Application PC/TUS0020131A
; GENERAL INFORMATION:
;  APPLICANT:  James S. HUSTON et al.
;  TITLE OF INVENTION:  METHODS AND COMPOSITIONS FOR INHIBITING POLYPEPTIDE
;  TITLE OF INVENTION:  ACCUMULATION ASSOCIATED WITH NEUROLOGICAL DISORDERS
;  FILE REFERENCE:  INR-004PC
;  CURRENT APPLICATION NUMBER: PCT/US00/20131A
;  CURRENT FILING DATE:  2000-07-24
;  PRIOR APPLICATION NUMBER:  60/146,047
;  PRIOR FILING DATE:  1999-07-27
;  NUMBER OF SEQ ID NOS:  45
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 43
;   LENGTH: 9
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: Synthetic
;   OTHER INFORMATION: construct
PCT-US00-20131A-43


  Query Match           100.0%;  Score 55;  DB 1;  Length 9;
  Best Local Similarity  100.0%;  Pred. No. 2.9e+06;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 YPYDVPDYA 9
          |||||||||
Db      1 YPYDVPDYA 9


RESULT   6
PCT-US00-22456-1
; Sequence 1, Application PC/TUS0022456
; GENERAL INFORMATION:
;  APPLICANT:  Donald J. Buchsbaum
;  APPLICANT:  Buck E. Rogers
;  APPLICANT:  Kurt R. Zinn
;  APPLICANT:  David T. Curiel
;  TITLE OF INVENTION:  Gene Transfer Imaging and Uses Thereof
;  FILE REFERENCE:  D5836CIPPCT
;  CURRENT APPLICATION NUMBER: PCT/US00/22456
;  CURRENT FILING DATE:  2000-08-16
;  PRIOR APPLICATION NUMBER:  09/374,972
;  PRIOR FILING DATE:  1999-08-16
;  NUMBER OF SEQ ID NOS:  6
; SEQ ID NO 1
```

```
;   LENGTH: 9
;   TYPE: PRT
;   ORGANISM: Artificial sequence
;   FEATURE:
;   OTHER INFORMATION: amino acid sequence of "HA" tag
PCT-US00-22456-1


  Query Match           100.0%;  Score 55;  DB 1;  Length 9;
  Best Local Similarity  100.0%;  Pred. No. 2.9e+06;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 YPYDVPDYA 9
          |||||||||
Db      1 YPYDVPDYA 9


RESULT   7
PCT-US00-23158-10
; Sequence 10, Application PC/TUS0023158
;   GENERAL INFORMATION:
;        APPLICANT: Cunningham, Sonia
;                   Barros, Maria Pia
;        TITLE OF INVENTION: A POLYNUCLEOTIDE ENCODING A HUMAN
;                            JUNCTIONAL ADHESION PROTEIN (JAM 2)
;        NUMBER OF SEQUENCES: 10
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Rockey, Milnamow & Katz, Ltd.
;             STREET: 180 N. Stetson Avenue, 2 Prudential Plaza,
;                     Suite 4700
;             CITY: Chicago
;             STATE: IL
;             COUNTRY: U.S.A.
;             ZIP: 60601
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.30
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: PCT/US00/23158
;             FILING DATE: 23-Aug-2000
;             CLASSIFICATION: <Unknown>
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Katz, Martin L.
;             REGISTRATION NUMBER: 25,011
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: 312-616-5400
;             TELEFAX: 312-616-5460
;   INFORMATION FOR SEQ ID NO: 10:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 9 amino acids
;             TYPE: amino acid
;             STRANDEDNESS: single
;             TOPOLOGY: linear
;        MOLECULE TYPE: peptide
;        SEQUENCE DESCRIPTION: SEQ ID NO: 10:
PCT-US00-23158-10


  Query Match           100.0%;  Score 55;  DB 1;  Length 9;
  Best Local Similarity  100.0%;  Pred. No. 2.9e+06;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 YPYDVPDYA 9
          |||||||||
Db      1 YPYDVPDYA 9


RESULT   8
PCT-US00-30238-5
; Sequence 5, Application PC/TUS0030238
; GENERAL INFORMATION:
;  APPLICANT:  Lok, Si
;  TITLE OF INVENTION:  Cell Surface Display of Proteins by
;  TITLE OF INVENTION:  Recombinant Host Cells
;  FILE REFERENCE:  99-34
;  CURRENT APPLICATION NUMBER: PCT/US00/30238
;  CURRENT FILING DATE:  2000-11-01
;  NUMBER OF SEQ ID NOS:  5
;  SOFTWARE:  FastSEQ for Windows Version 3.0
; SEQ ID NO 5
;   LENGTH: 9
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Hemagglutinin A epitope tag
PCT-US00-30238-5


  Query Match           100.0%;  Score 55;  DB 1;  Length 9;
  Best Local Similarity  100.0%;  Pred. No. 2.9e+06;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 YPYDVPDYA 9
          |||||||||
Db      1 YPYDVPDYA 9


RESULT   9
PCT-US93-12643-6
; Sequence 6, Application PC/TUS9312643
;  GENERAL INFORMATION:
;    APPLICANT:  Roger Brent
;    APPLICANT:  Antonis S. Zervos
;    TITLE OF INVENTION:  MAX-INTERACTING PROTEINS AND
;    TITLE OF INVENTION:  RELATED MOLECULES AND METHODS
;    NUMBER OF SEQUENCES:  11
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Fish & Richardson
;      STREET:  225 Franklin Street
;      CITY:  Boston
;      STATE:  Massachusetts
;      COUNTRY:  U.S.A.
;      ZIP:  02110-2804
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5" Diskette, 1.44 Mb
;      COMPUTER:  IBM PS/2 Model 50Z or 55SX
;      OPERATING SYSTEM:  MS-DOS (Version 5.0)
;      SOFTWARE:  WordPerfect (Version 5.1)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US93/12643
;      FILING DATE:  29-DEC-1993
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/011,398
;      FILING DATE:  29-JAN-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Paul T. Clark
;      REGISTRATION NUMBER:  30,162
;      REFERENCE/DOCKET NUMBER:  00786/160001
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617) 542-5070
;      TELEFAX:  (617) 542-8906
;      TELEX:  200154
;  INFORMATION FOR SEQ ID NO:  6:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  9
;      TYPE:  amino acid
;      STRANDEDNESS:  N/A
;      TOPOLOGY:  N/A
PCT-US93-12643-6


  Query Match           100.0%;  Score 55;  DB 1;  Length 9;
  Best Local Similarity  100.0%;  Pred. No. 2.9e+06;
```

```
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 YPYDVPDYA 9
          |||||||||
Db      1 YPYDVPDYA 9


RESULT  10
PCT-US95-01467-14
; Sequence 14, Application PC/TUS9501467
;  GENERAL INFORMATION:
;    APPLICANT:  THE SCRIPPS RESEARCH INSTITUTE
;    TITLE OF INVENTION:  METHOD FOR USING TOBACCO MOSAIC VIRUS TO
;    TITLE OF INVENTION:  OVERPRODUCE PEPTIDES AND PROTEINS
;    NUMBER OF SEQUENCES:  30
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Spensley Horn Jubas & Lubitz
;      STREET:  1880 Century Park East, Suite 500
;      CITY:  Los Angeles
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  90067
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US95/01467
;      FILING DATE:  03-FEB-1995
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Bostich, June M.
;      REGISTRATION NUMBER:  31,238
;      REFERENCE/DOCKET NUMBER:  FD-4074
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (619) 455-5100
;      TELEFAX:  (619) 455-5110
;  INFORMATION FOR SEQ ID NO:  14:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  9 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    IMMEDIATE SOURCE:
;      CLONE:  12CA5
;    FEATURE:
;      NAME/KEY:  Peptide
;      LOCATION:  1..9
PCT-US95-01467-14


  Query Match           100.0%;  Score 55;  DB 1;  Length 9;
  Best Local Similarity  100.0%;  Pred. No. 2.9e+06;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 YPYDVPDYA 9
          |||||||||
Db      1 YPYDVPDYA 9


RESULT  11
PCT-US97-15500-12
; Sequence 12, Application PC/TUS9715500
;  GENERAL INFORMATION:
;    APPLICANT:  The General Hospital Corporation
;    TITLE OF INVENTION:  INHIBITION OF VIRAL REPLICATION
;    NUMBER OF SEQUENCES:  21
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Fish & Richardson P.C.
;      STREET:  225 Franklin Street
;      CITY:  Boston
;      STATE:  MA
;      COUNTRY:  USA
;      ZIP:  02110-2804
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  Windows
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US97/15500
;      FILING DATE:  03-SEP-1997
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  60/025,370
;      FILING DATE:  03-SEP-1996
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Fraser, Janis K.
;      REGISTRATION NUMBER:  34,819
;      REFERENCE/DOCKET NUMBER:  08472/705W01
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617/542-5070
;      TELEFAX:  617/542-8906
;      TELEX:  200154
;  INFORMATION FOR SEQ ID NO:  12:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  9 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  not relevant
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
PCT-US97-15500-12


  Query Match           100.0%;  Score 55;  DB 1;  Length 9;
  Best Local Similarity  100.0%;  Pred. No. 2.9e+06;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 YPYDVPDYA 9
          |||||||||
Db      1 YPYDVPDYA 9


RESULT  12
PCT-US97-20344-11
; Sequence 11, Application PC/TUS9720344
;  GENERAL INFORMATION:
;    APPLICANT:
;    TITLE OF INVENTION:  COMPOSITIONS AND USES FOR A NOVEL CELL
;    TITLE OF INVENTION:  DEATH PROTECTING PROTEIN
;    NUMBER OF SEQUENCES:  12
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30 (EPO)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US97/20344
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US Unknown
;      FILING DATE:  04-NOV-1997
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/030,302
;      FILING DATE:  05-NOV-1996
;  INFORMATION FOR SEQ ID NO:  11:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  9 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  linear
PCT-US97-20344-11
```

```
  Query Match           100.0%;  Score 55;  DB 1;  Length 9;
  Best Local Similarity  100.0%;  Pred. No. 2.9e+06;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 YPYDVPDYA 9
          |||||||||
Db      1 YPYDVPDYA 9


RESULT  13
PCT-US97-22198-1
; Sequence 1, Application PC/TUS9722198
;  GENERAL INFORMATION:
;    APPLICANT:
;    TITLE OF INVENTION:  INACTIVATION OF HIV CO-RECEPTORS AS THERAPY
;    TITLE OF INVENTION:  FOR HIV INFECTION
;    NUMBER OF SEQUENCES:  6
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30 (EPO)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US97/22198
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/032,277
;      FILING DATE:  02-DEC-1996
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  9 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  linear
PCT-US97-22198-1


  Query Match           100.0%;  Score 55;  DB 1;  Length 9;
  Best Local Similarity  100.0%;  Pred. No. 2.9e+06;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 YPYDVPDYA 9
          |||||||||
Db      1 YPYDVPDYA 9


RESULT  14
PCT-US98-06791-1
; Sequence 1, Application PC/TUS9806791
;  GENERAL INFORMATION:
;    APPLICANT:  PORT, J. D.
;    APPLICANT:  BRISTOW, MICHAEL R.
;    TITLE OF INVENTION:  "TRANSGENIC MODEL AND TREATMENT FOR
;    TITLE OF INVENTION:  HEART DISEASE"
;    NUMBER OF SEQUENCES:  5
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  SHERIDAN ROSS P.C.
;      STREET:  1700 LINCOLN ST., SUITE 3500
;      CITY:  DENVER
;      STATE:  COLORADO
;      COUNTRY:  USA
;      ZIP:  80203
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US98/06791
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/041,966
;      FILING DATE:  03-AUG-1997
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  CROOK, WANNELL M.
;      REGISTRATION NUMBER:  31,071
;      REFERENCE/DOCKET NUMBER:  3595-2-PCT
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (303) 863-9700
;      TELEFAX:  (303) 863-0223
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  9 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
PCT-US98-06791-1


  Query Match           100.0%;  Score 55;  DB 1;  Length 9;
  Best Local Similarity  100.0%;  Pred. No. 2.9e+06;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 YPYDVPDYA 9
          |||||||||
Db      1 YPYDVPDYA 9


RESULT  15
PCT-US98-11189-28
; Sequence 28, Application PC/TUS9811189
;  GENERAL INFORMATION:
;    APPLICANT:  Nimni, Marcel E.
;    APPLICANT:  Hall, Frederick L.
;    APPLICANT:  Wu, Lingtao
;    APPLICANT:  Han, Bo
;    APPLICANT:  Shors, Edwin
;    TITLE OF INVENTION:  Bone Morphogenetic
;    TITLE OF INVENTION:  Proteins and Their Use in Bone Growth
;    NUMBER OF SEQUENCES:  46
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Merchant & Gould
;      STREET:  11150 Santa Monica Boulevard, Suite 400
;      CITY:  Los Angeles
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  90025-3395
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  FastSeq Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US98/11189
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/868,452
;      FILING DATE:  3-June-1997
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Wood, William J.
;      REGISTRATION NUMBER:  P-42,236
;      REFERENCE/DOCKET NUMBER:  30853-1WO01
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  310-445-1140
;      TELEFAX:  310-445-9031
;  INFORMATION FOR SEQ ID NO:  28:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  9 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
```

; MOLECULE TYPE: protein
PCT-US98-11189-28

Query Match 100.0%; Score 55; DB 1; Length 9;
Best Local Similarity 100.0%; Pred. No. 2.9e+06;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy      1 YPYDVPDYA 9
          |||||||||
Db      1 YPYDVPDYA 9
```

Search completed: April 11, 2002, 10:20:50
Job time: 400 sec

GenCore version 4.5

OM protein - protein search, using sw model

Run on: April 11, 2002, 10:14:10 ; Search time 39.37 Seconds
(without alignments)
16.824 Million cell updates/sec

Title: US-09-284-787-1
Perfect score: 55
Sequence: 1 YPYDVPDYA 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 367557 seqs, 73597022 residues

Total number of hits satisfying chosen parameters: 367557

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending_Patents_AA_New:*
1: /cgn2_6/ptodata/2/paa/PCT_NEW_COMB.pep:*
2: /cgn2_6/ptodata/2/paa/US06_NEW_COMB.pep:*
3: /cgn2_6/ptodata/2/paa/US07_NEW_COMB.pep:*
4: /cgn2_6/ptodata/2/paa/US08_NEW_COMB.pep:*
5: /cgn2_6/ptodata/2/paa/US09_NEW_COMB.pep:*
6: /cgn2_6/ptodata/2/paa/US09_NEW_COMB.pep1:*
7: /cgn2_6/ptodata/2/paa/US10_NEW_COMB.pep:*
8: /cgn2_6/ptodata/2/paa/US60_NEW_COMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	55	100.0	9	1	PCT-US01-45076-9	Sequence 9, Appli
2	55	100.0	9	5	US-09-466-568-19	Sequence 19, Appl
3	55	100.0	9	5	US-09-931-381A-24	Sequence 24, Appl
4	55	100.0	9	5	US-09-936-969-9	Sequence 9, Appli
5	55	100.0	9	5	US-09-180-167A-17	Sequence 17, Appl
6	55	100.0	9	5	US-09-775-052A-17	Sequence 17, Appl
7	55	100.0	9	6	US-09-912-733-32	Sequence 32, Appl
8	55	100.0	9	6	US-09-574-735C-45	Sequence 45, Appl
9	55	100.0	9	6	US-09-852-370A-33	Sequence 33, Appl
10	55	100.0	9	6	US-09-906-393A-4	Sequence 4, Appli
11	55	100.0	9	6	US-09-775-052-17	Sequence 17, Appl
12	55	100.0	9	6	US-09-554-726A-53	Sequence 53, Appl
13	55	100.0	9	6	US-09-845-667-32	Sequence 32, Appl
14	55	100.0	9	6	US-09-995-847-4	Sequence 4, Appli
15	55	100.0	9	6	US-09-554-726A-53	Sequence 53, Appl
16	55	100.0	9	6	US-09-529-106A-16	Sequence 16, Appl
17	55	100.0	9	6	US-09-713-858-7	Sequence 7, Appli
18	55	100.0	9	6	US-09-714-471-7	Sequence 7, Appli
19	55	100.0	9	6	US-09-896-915-37	Sequence 37, Appl
20	55	100.0	9	6	US-09-598-347-7	Sequence 7, Appli
21	55	100.0	9	6	US-09-830-980-8	Sequence 8, Appli
22	55	100.0	9	6	US-09-848-813-5	Sequence 5, Appli
23	55	100.0	9	6	US-09-918-036-17	Sequence 17, Appl
24	55	100.0	9	6	US-09-922-226-40	Sequence 40, Appl
25	55	100.0	9	6	US-09-807-305A-3	Sequence 3, Appli
26	55	100.0	9	6	US-09-352-171-7	Sequence 7, Appli
27	55	100.0	9	6	US-09-620-955B-28	Sequence 28, Appl
28	55	100.0	9	6	US-09-620-955B-43	Sequence 43, Appl
29	55	100.0	9	7	US-10-001-073-23	Sequence 23, Appl
30	55	100.0	9	7	US-10-066-151-93	Sequence 93, Appl
31	55	100.0	9	7	US-10-003-496-16	Sequence 16, Appl
32	55	100.0	9	7	US-10-009-178-10	Sequence 10, Appl
33	55	100.0	9	7	US-10-083-815-7	Sequence 7, Appli
34	55	100.0	9	7	US-10-059-720-49	Sequence 49, Appl
35	55	100.0	9	7	US-10-084-706-55	Sequence 55, Appl
36	55	100.0	10	5	US-09-563-055A-1	Sequence 1, Appli
37	55	100.0	10	6	US-09-453-234-30	Sequence 30, Appl
38	55	100.0	10	6	US-09-589-619-10	Sequence 10, Appl
39	55	100.0	10	6	US-09-563-055-1	Sequence 1, Appli
40	55	100.0	10	6	US-09-994-487-1	Sequence 1, Appli
41	55	100.0	10	6	US-09-645-456-25	Sequence 25, Appl
42	55	100.0	10	6	US-09-900-590A-28	Sequence 28, Appl
43	55	100.0	10	6	US-09-822-295-13	Sequence 13, Appl
44	55	100.0	11	5	US-09-415-765A-2	Sequence 2, Appli
45	55	100.0	11	5	US-09-626-580B-2	Sequence 2, Appli

ALIGNMENTS

```
RESULT   1
PCT-US01-45076-9
; Sequence 9, Application PC/TUS0145076
; GENERAL INFORMATION:
;  APPLICANT: The General Hospital Corporation
;  APPLICANT:  Scadden, David T.
;  APPLICANT:  Lee, Byeong-Chel
;  TITLE OF INVENTION: METHODS AND PRODUCTS FOR MANIPULATING
;  TITLE OF INVENTION:  HEMATOPOIETIC STEM CELLS
;  FILE REFERENCE: M0765/7039WO/ERG/KA
;  CURRENT APPLICATION NUMBER: PCT/US01/45076
;  CURRENT FILING DATE:  2001-11-29
;  PRIOR APPLICATION NUMBER: U.S. 60/250,727
;  PRIOR FILING DATE: 2000-12-01
;  NUMBER OF SEQ ID NOS: 10
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 9
;   LENGTH: 9
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: HA-tag
PCT-US01-45076-9

  Query Match             100.0%;  Score 55;  DB 1;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 3.1e+05;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 YPYDVPDYA 9
           |||||||||
Db       1 YPYDVPDYA 9


RESULT   2
US-09-466-568-19
; Sequence 19, Application US/09466568
;   GENERAL INFORMATION:
;        APPLICANT: Crabtree, Gerald R.
;                   Schreiber, Stuart L.
;                   Spencer, David M.
;                   Wandless, Thomas J.
;                   Belshaw, Peter
;        TITLE OF INVENTION: REGULATED TRANSCRIPTION OF TARGETED
;                            GENES AND OTHER BIOLOGICAL EVENTS
;        NUMBER OF SEQUENCES: 81
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: ARIAD Pharmaceuticals, Inc.
```

```
;             STREET: 26 Landsdowne Street
;             CITY: Cambridge
;             STATE: Massachusetts
;             COUNTRY: USA
;             ZIP: 02139
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC/DOS/MS/DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.25
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/466,568
;             FILING DATE: 17-Dec-1999
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: US/09/157,753
;             FILING DATE: <Unknown>
;             APPLICATION NUMBER: 08/388,653
;             FILING DATE: 14-FEB-1995
;             APPLICATION NUMBER: US 08/478,386
;             FILING DATE: 07-JUN-1995
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Figg, E. Anthony
;             REGISTRATION NUMBER: 27,195
;             REFERENCE/DOCKET NUMBER: 2054-114A
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (202) 783-6040
;             TELEFAX: (202) 783-6031
;   INFORMATION FOR SEQ ID NO: 19:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 9 amino acids
;             TYPE: amino acid
;             TOPOLOGY: linear
;        MOLECULE TYPE: protein
;        SEQUENCE DESCRIPTION: SEQ ID NO: 19:
US-09-466-568-19


  Query Match           100.0%;  Score 55;  DB 5;  Length 9;
  Best Local Similarity  100.0%;  Pred. No. 3.1e+05;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 YPYDVPDYA 9
          |||||||||
Db      1 YPYDVPDYA 9


RESULT   3
US-09-931-381A-24
; Sequence 24, Application US/09931381A
; GENERAL INFORMATION:
;  APPLICANT: Butcher, Eugene C.
;  APPLICANT:  Kunkel, Eric J.
;  APPLICANT:  Pan, Junliang
;  APPLICANT:  Soler-Ferran, Dulce
;  TITLE OF INVENTION: Method for Identifying Agents Which
;  TITLE OF INVENTION:  Modulate Chemokine "MEC"-Induced Functions of CCR3 and/or
;  TITLE OF INVENTION:  CCR10
;  FILE REFERENCE: 1855.2010-003
;  CURRENT APPLICATION NUMBER: US/09/931,381A
;  CURRENT FILING DATE:  2001-08-15
;  PRIOR APPLICATION NUMBER: U.S. 09/638,914
;  PRIOR FILING DATE: 2000-08-15
;  NUMBER OF SEQ ID NOS: 24
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 24
;   LENGTH: 9
;   TYPE: PRT
;   ORGANISM: Influenza virus
US-09-931-381A-24


  Query Match           100.0%;  Score 55;  DB 5;  Length 9;
  Best Local Similarity  100.0%;  Pred. No. 3.1e+05;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 YPYDVPDYA 9
          |||||||||
Db      1 YPYDVPDYA 9


RESULT   4
US-09-936-969-9
; Sequence 9, Application US/09936969
; GENERAL INFORMATION:
;  APPLICANT: Iris, F.
;  APPLICANT:  Pourny, J.
;  TITLE OF INVENTION: METHODS FOR DETECTION OF NUCLEIC ACID POLYMORPHISMS USING
;  TITLE OF INVENTION:  PEPTIDE-LABELED OLIGONUCLEOTIDES AND ANTIBODY ARRAYS
;  FILE REFERENCE: 9408-042-999
;  CURRENT APPLICATION NUMBER: US/09/936,969
;  CURRENT FILING DATE:  2001-09-18
;  NUMBER OF SEQ ID NOS: 11
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 9
;   LENGTH: 9
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Oryctolagus cuniculus
US-09-936-969-9


  Query Match           100.0%;  Score 55;  DB 5;  Length 9;
  Best Local Similarity  100.0%;  Pred. No. 3.1e+05;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 YPYDVPDYA 9
          |||||||||
Db      1 YPYDVPDYA 9


RESULT   5
US-09-180-167A-17
; Sequence 17, Application US/09180167A
; GENERAL INFORMATION:
;  APPLICANT: Gordon C. Shore et al.
;  TITLE OF INVENTION: METHODS AND REAGENTS FOR MODULATING
;  TITLE OF INVENTION:  APOPTOSIS
;  FILE REFERENCE: 50013/004003
;  CURRENT APPLICATION NUMBER: US/09/180,167A
;  CURRENT FILING DATE:  1999-09-10
;  PRIOR APPLICATION NUMBER: CA 2,198,988
;  PRIOR FILING DATE: 1997-03-03
;  PRIOR APPLICATION NUMBER: PCT/IB98/00706
;  PRIOR FILING DATE: 1998-03-02
;  NUMBER OF SEQ ID NOS: 33
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 17
;   LENGTH: 9
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Synthetic, based on Orthomyxoviridae Influenza A
;   OTHER INFORMATION: virus
US-09-180-167A-17


  Query Match           100.0%;  Score 55;  DB 5;  Length 9;
  Best Local Similarity  100.0%;  Pred. No. 3.1e+05;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 YPYDVPDYA 9
          |||||||||
Db      1 YPYDVPDYA 9
```

```
RESULT   6
US-09-775-052A-17
; Sequence 17, Application US/09775052A
; GENERAL INFORMATION:
;  APPLICANT: Dowdy, Steven F.
;  TITLE OF INVENTION: ANTI-PATHOGEN SYSTEM AND METHODS OF USE THEREOF
;  FILE REFERENCE: 48881/1742
;  CURRENT APPLICATION NUMBER: US/09/775,052A
;  CURRENT FILING DATE:  2001-12-05
;  PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US/09/208,966
;  PRIOR FILING DATE: EARLIER FILING DATE: 1998-12-10
;  PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/069,012
;  PRIOR FILING DATE: EARLIER FILING DATE: 1997-12-10
;  NUMBER OF SEQ ID NOS: 57
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 17
;   LENGTH: 9
;   TYPE: PRT
;   ORGANISM: human
US-09-775-052A-17


  Query Match           100.0%;  Score 55;  DB 5;  Length 9;
  Best Local Similarity  100.0%;  Pred. No. 3.1e+05;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 YPYDVPDYA 9
          |||||||||
Db      1 YPYDVPDYA 9


RESULT   7
US-09-912-733-32
; Sequence 32, Application US/09912733
; GENERAL INFORMATION:
;  APPLICANT: Ward, Michael
;  APPLICANT:  Power, Scott D.
;  TITLE OF INVENTION: DNA SEQUENCES, VECTORS, AND FUSION POLYPEPTIDES FOR SECRETION OF
;  TITLE OF INVENTION:  POLYPEPTIDES IN FILAMENTOUS FUNGI
;  FILE REFERENCE: A-62926-1
;  CURRENT APPLICATION NUMBER: US/09/912,733
;  CURRENT FILING DATE:  2001-07-23
;  PRIOR APPLICATION NUMBER: US 08/785,668
;  PRIOR FILING DATE: 1997-01-17
;  NUMBER OF SEQ ID NOS: 44
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 32
;   LENGTH: 9
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Synthetic.
US-09-912-733-32


  Query Match           100.0%;  Score 55;  DB 6;  Length 9;
  Best Local Similarity  100.0%;  Pred. No. 3.1e+05;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 YPYDVPDYA 9
          |||||||||
Db      1 YPYDVPDYA 9


RESULT   8
US-09-574-735C-45
; Sequence 45, Application US/09574735C
; GENERAL INFORMATION:
;  APPLICANT: De Veylder, Lieven
;  APPLICANT:  Beeckman, Tom
;  APPLICANT:  Inz,, Dirk
;  APPLICANT:  Van Camp, Wim
;  APPLICANT:  Krols, Luc
;  TITLE OF INVENTION: Cyclin-dependent kinase inhibitors and uses thereof
;  FILE REFERENCE: 2283/301
;  CURRENT APPLICATION NUMBER: US/09/574,735C
;  CURRENT FILING DATE:  2000-05-18
;  NUMBER OF SEQ ID NOS: 48
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 45
;   LENGTH: 9
;   TYPE: PRT
;   ORGANISM: HA-epitope
US-09-574-735C-45


  Query Match           100.0%;  Score 55;  DB 6;  Length 9;
  Best Local Similarity  100.0%;  Pred. No. 3.1e+05;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 YPYDVPDYA 9
          |||||||||
Db      1 YPYDVPDYA 9


RESULT   9
US-09-852-370A-33
; Sequence 33, Application US/09852370A
; GENERAL INFORMATION:
;  APPLICANT: Pomerantz, Joel L.
;  APPLICANT:  Sharp, Phillip A.
;  APPLICANT:  Pabo, Carl O.
;  TITLE OF INVENTION: Chimeric DNA-binding proteins
;  FILE REFERENCE: APV-022.02
;  CURRENT APPLICATION NUMBER: US/09/852,370A
;  CURRENT FILING DATE:  2001-05-10
;  PRIOR APPLICATION NUMBER: 08/973,131
;  PRIOR FILING DATE: 1997-11-26
;  PRIOR APPLICATION NUMBER: PCT/US95/16982
;  PRIOR FILING DATE: 1995-12-29
;  PRIOR APPLICATION NUMBER: 08/366,083
;  PRIOR FILING DATE: 1994-12-29
;  NUMBER OF SEQ ID NOS: 75
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 33
;   LENGTH: 9
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: encoded
;   OTHER INFORMATION: epitope
US-09-852-370A-33


  Query Match           100.0%;  Score 55;  DB 6;  Length 9;
  Best Local Similarity  100.0%;  Pred. No. 3.1e+05;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 YPYDVPDYA 9
          |||||||||
Db      1 YPYDVPDYA 9


RESULT  10
US-09-906-393A-4
; Sequence 4, Application US/09906393A
; GENERAL INFORMATION:
;  APPLICANT: Wang, Zhou
;  APPLICANT:  Xiao, Wuhan
;  TITLE OF INVENTION: METHOD OF PROGNOSING CANCER AND THE PROTEINS INVOLVED
;  FILE REFERENCE: 1720-1-001CIP
;  CURRENT APPLICATION NUMBER: US/09/906,393A
;  CURRENT FILING DATE:  2001-07-16
;  PRIOR APPLICATION NUMBER: 60/218,761
;  PRIOR FILING DATE: 2000-07-17
```

```
;  NUMBER OF SEQ ID NOS: 36
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
;   LENGTH: 9
;   TYPE: PRT
;   ORGANISM: Artificial sequence
;   FEATURE:
;   OTHER INFORMATION: HA Tag
US-09-906-393A-4


  Query Match           100.0%;  Score 55;  DB 6;  Length 9;
  Best Local Similarity  100.0%;  Pred. No. 3.1e+05;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 YPYDVPDYA 9
          |||||||||
Db      1 YPYDVPDYA 9


RESULT  11
US-09-775-052-17
; Sequence 17, Application US/09775052
; GENERAL INFORMATION:
;  APPLICANT: Dowdy, Steven F.
;  TITLE OF INVENTION: ANTI-PATHOGEN SYSTEM AND METHODS OF USE THEREOF
;  FILE REFERENCE: 48881/1742
;  CURRENT APPLICATION NUMBER: US/09/775,052
;  CURRENT FILING DATE:  2001-02-01
;  PRIOR APPLICATION NUMBER: 09/208,966
;  PRIOR FILING DATE: 1998-12-10
;  PRIOR APPLICATION NUMBER: 60/082,402
;  PRIOR FILING DATE: 1998-04-20
;  PRIOR APPLICATION NUMBER: 60/069,012
;  PRIOR FILING DATE: 1997-12-10
;  NUMBER OF SEQ ID NOS: 57
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 17
;   LENGTH: 9
;   TYPE: PRT
;   ORGANISM: human
US-09-775-052-17


  Query Match           100.0%;  Score 55;  DB 6;  Length 9;
  Best Local Similarity  100.0%;  Pred. No. 3.1e+05;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 YPYDVPDYA 9
          |||||||||
Db      1 YPYDVPDYA 9


RESULT  12
US-09-554-726A-53
; Sequence 53, Application US/09554726A
; GENERAL INFORMATION:
;  APPLICANT: HERRMANN, Bernhard
;  APPLICANT:  KOSCHORZ, Birgit
;  APPLICANT:  KISPERT, Andreas
;  TITLE OF INVENTION: NUCLEIC ACIDS INVOLVED IN THE RESPONDER PHENOTYPE AND APPLICATION
;  TITLE OF INVENTION:  THEREOF
;  FILE REFERENCE: 258.0009 0101
;  CURRENT APPLICATION NUMBER: US/09/554,726A
;  CURRENT FILING DATE:  2000-05-18
;  PRIOR APPLICATION NUMBER: PCT/EP 98/07395
;  PRIOR FILING DATE: 1998-11-18
;  PRIOR APPLICATION NUMBER: EP 98 10 3596.7
;  PRIOR FILING DATE: 1998-03-02
;  PRIOR APPLICATION NUMBER: EP 97 12 0190.0
;  PRIOR FILING DATE: 1997-11-18
;  NUMBER OF SEQ ID NOS: 53
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 53
;   LENGTH: 9
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Peptide
US-09-554-726A-53


  Query Match           100.0%;  Score 55;  DB 6;  Length 9;
  Best Local Similarity  100.0%;  Pred. No. 3.1e+05;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 YPYDVPDYA 9
          |||||||||
Db      1 YPYDVPDYA 9


RESULT  13
US-09-845-667-32
; Sequence 32, Application US/09845667
;   GENERAL INFORMATION:
;        APPLICANT: Cohen, Philip
;                   Alessi, Dario
;                   Cross, Darren
;        TITLE OF INVENTION: CONTROL OF PROTEIN SYNTHESIS, AND SCREENING METHOD
;                            FOR AGENTS
;        NUMBER OF SEQUENCES: 58
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Braman & Rogalskyj, LLP
;             STREET: P.O. Box 352
;             CITY: Canandaigua
;             STATE: New York
;             COUNTRY: USA
;             ZIP: 14424-0352
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.30
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/845,667
;             FILING DATE: 30-Apr-2001
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: US 09/091,763
;             FILING DATE: 19-JUN-1998
;             APPLICATION NUMBER: PCT/GB96/03186
;             FILING DATE: 20-DEC-1996
;             APPLICATION NUMBER: GB 9526083.2
;             FILING DATE: 20-DEC-1995
;             APPLICATION NUMBER: GB 9610272.8
;             FILING DATE: 16-MAY-1996
;             APPLICATION NUMBER: GB 9615066.9
;             FILING DATE: 18-JUL-1996
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Braman, Susan J
;             REGISTRATION NUMBER: 34,103
;             REFERENCE/DOCKET NUMBER: 002.00041
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: 716-393-3002
;             TELEFAX: 716-393-3001
;   INFORMATION FOR SEQ ID NO: 32:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 9 amino acids
;             TYPE: amino acid
;             STRANDEDNESS: single
;             TOPOLOGY: linear
;        MOLECULE TYPE: peptide
;        SEQUENCE DESCRIPTION: SEQ ID NO: 32:
US-09-845-667-32
```

```
  Query Match           100.0%;  Score 55;  DB 6;  Length 9;
  Best Local Similarity  100.0%;  Pred. No. 3.1e+05;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 YPYDVPDYA 9
          |||||||||
Db      1 YPYDVPDYA 9


RESULT  14
US-09-995-847-4
; Sequence 4, Application US/09995847
; GENERAL INFORMATION:
;  APPLICANT: Rizzuto, Carlo
;  APPLICANT:  Afeyan, Noubar
;  APPLICANT:  Lee, Frank
;  APPLICANT:  Church, George
;  APPLICANT:  Das Gupta, Ruchira
;  APPLICANT:  Zhang, Bin
;  APPLICANT:  Schwartz, John
;  APPLICANT:  Lugovskoy, Alexey
;  TITLE OF INVENTION: MODULAR MOLECULAR CLASPS AND USES THEREOF
;  FILE REFERENCE: ENZ-001
;  CURRENT APPLICATION NUMBER: US/09/995,847
;  CURRENT FILING DATE:  2001-11-28
;  PRIOR APPLICATION NUMBER: 60/279,524
;  PRIOR FILING DATE: 2001-03-28
;  NUMBER OF SEQ ID NOS: 26
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
;   LENGTH: 9
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:HA epitope
US-09-995-847-4


  Query Match           100.0%;  Score 55;  DB 6;  Length 9;
  Best Local Similarity  100.0%;  Pred. No. 3.1e+05;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 YPYDVPDYA 9
          |||||||||
Db      1 YPYDVPDYA 9


RESULT  15
US-09-554-726A-53
; Sequence 53, Application US/09554726A
; GENERAL INFORMATION:
;  APPLICANT: HERRMANN, Bernhard
;  APPLICANT:  KOSCHORZ, Birgit
;  APPLICANT:  KISPERT, Andreas
;  TITLE OF INVENTION: NUCLEIC ACIDS INVOLVED IN THE RESPONDER PHENOTYPE AND APPLICATION
;  TITLE OF INVENTION:  THEREOF
;  FILE REFERENCE: 258.0009 0101
;  CURRENT APPLICATION NUMBER: US/09/554,726A
;  CURRENT FILING DATE:  2000-05-18
;  PRIOR APPLICATION NUMBER: PCT/EP 98/07395
;  PRIOR FILING DATE: 1998-11-18
;  PRIOR APPLICATION NUMBER: EP 98 10 3596.7
;  PRIOR FILING DATE: 1998-03-02
;  PRIOR APPLICATION NUMBER: EP 97 12 0190.0
;  PRIOR FILING DATE: 1997-11-18
;  NUMBER OF SEQ ID NOS: 53
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 53
;   LENGTH: 9
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Peptide
US-09-554-726A-53


  Query Match           100.0%;  Score 55;  DB 6;  Length 9;
  Best Local Similarity  100.0%;  Pred. No. 3.1e+05;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 YPYDVPDYA 9
          |||||||||
Db      1 YPYDVPDYA 9


Search completed: April 11, 2002, 10:21:36
Job time: 446 sec
```

THIS PAGE BLANK (USPTO)

GenCore version 4.5

OM protein - protein search, using sw model

Run on: April 11, 2002, 10:20:50 ; Search time 266.73 Seconds
(without alignments)
14.574 Million cell updates/sec

Title: US-09-284-787-2
Perfect score: 80
Sequence: 1 YPYDVPDYAGSGSK 14

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 3148936 seqs, 277657034 residues

Total number of hits satisfying chosen parameters: 3148936

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending_Patents_AA_Main:*
1: /cgn2_6/ptodata/2/paa/PCTUS_COMB.pep:*
2: /cgn2_6/ptodata/2/paa/US06_COMB.pep:*
3: /cgn2_6/ptodata/2/paa/US07_COMB.pep:*
4: /cgn2_6/ptodata/2/paa/US080_COMB.pep:*
5: /cgn2_6/ptodata/2/paa/US081_COMB.pep:*
6: /cgn2_6/ptodata/2/paa/US082_COMB.pep:*
7: /cgn2_6/ptodata/2/paa/US083_COMB.pep:*
8: /cgn2_6/ptodata/2/paa/US084_COMB.pep:*
9: /cgn2_6/ptodata/2/paa/US085_COMB.pep:*
10: /cgn2_6/ptodata/2/paa/US086_COMB.pep:*
11: /cgn2_6/ptodata/2/paa/US087_COMB.pep:*
12: /cgn2_6/ptodata/2/paa/US088_COMB.pep:*
13: /cgn2_6/ptodata/2/paa/US089_COMB.pep:*
14: /cgn2_6/ptodata/2/paa/US090_COMB.pep:*
15: /cgn2_6/ptodata/2/paa/US091_COMB.pep:*
16: /cgn2_6/ptodata/2/paa/US092_COMB.pep:*
17: /cgn2_6/ptodata/2/paa/US093_COMB.pep:*
18: /cgn2_6/ptodata/2/paa/US094_COMB.pep:*
19: /cgn2_6/ptodata/2/paa/US095_COMB.pep:*
20: /cgn2_6/ptodata/2/paa/US096_COMB.pep:*
21: /cgn2_6/ptodata/2/paa/US097_COMB.pep:*
22: /cgn2_6/ptodata/2/paa/US098_COMB.pep:*
23: /cgn2_6/ptodata/2/paa/US099_COMB.pep:*
24: /cgn2_6/ptodata/2/paa/US60_COMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	80	100.0	14	16	US-09-284-787-2	Sequence 2, Appli
2	75	93.8	165	21	US-09-780-224A-11	Sequence 11, Appl
3	68	85.0	32	18	US-09-475-158-56	Sequence 56, Appl
4	67	83.8	126	20	US-09-622-646-23	Sequence 23, Appl
5	67	83.8	143	20	US-09-622-646-22	Sequence 22, Appl
6	67	83.8	147	20	US-09-622-646-25	Sequence 25, Appl
7	65	81.2	16	1	PCT-US00-09002-11	Sequence 11, Appl
8	65	81.2	16	1	PCT-US98-16887-7	Sequence 7, Appli
9	65	81.2	16	15	US-09-134-793-7	Sequence 7, Appli
10	65	81.2	16	15	US-09-134-793A-7	Sequence 7, Appli
11	65	81.2	16	17	US-09-300-639B-4	Sequence 4, Appli
12	65	81.2	16	18	US-09-459-113A-15	Sequence 15, Appl
13	65	81.2	16	19	US-09-514-673-15	Sequence 15, Appl
14	65	81.2	17	11	US-08-787-738-51	Sequence 51, Appl
15	65	81.2	17	11	US-08-787-738B-51	Sequence 51, Appl
16	65	81.2	17	13	US-08-963-368-51	Sequence 51, Appl
17	65	81.2	25	10	US-08-642-120-27	Sequence 27, Appl
18	65	81.2	28	5	US-08-108-351-4	Sequence 4, Appli
19	65	81.2	346	16	US-09-215-569B-8	Sequence 8, Appli
20	63	78.8	15	13	US-08-935-088-5	Sequence 5, Appli
21	62	77.5	219	16	US-09-201-931-2	Sequence 2, Appli
22	62	77.5	219	22	US-09-847-185-2	Sequence 2, Appli
23	61	76.2	10	19	US-09-546-238-13	Sequence 13, Appl
24	61	76.2	10	23	US-09-954-043-13	Sequence 13, Appl
25	61	76.2	12	14	US-09-023-905-16	Sequence 16, Appl
26	61	76.2	21	7	US-08-361-811-127	Sequence 127, App
27	61	76.2	21	7	US-08-381-478-169	Sequence 169, App
28	61	76.2	21	9	US-08-535-025-169	Sequence 169, App
29	61	76.2	21	9	US-08-591-919-169	Sequence 169, App
30	61	76.2	21	9	US-08-592-007-169	Sequence 169, App
31	61	76.2	21	9	US-08-592-027-169	Sequence 169, App
32	61	76.2	21	9	US-08-592-131-169	Sequence 169, App
33	61	76.2	21	13	US-08-964-039-169	Sequence 169, App
34	61	76.2	45	23	US-09-921-144-6	Sequence 6, Appli
35	61	76.2	48	23	US-09-921-144-8	Sequence 8, Appli
36	61	76.2	67	23	US-09-921-144-10	Sequence 10, Appl
37	61	76.2	70	23	US-09-921-144-12	Sequence 12, Appl
38	61	76.2	162	23	US-09-921-144-16	Sequence 16, Appl
39	61	76.2	163	23	US-09-921-144-14	Sequence 14, Appl
40	61	76.2	184	23	US-09-921-144-20	Sequence 20, Appl
41	61	76.2	185	23	US-09-921-144-18	Sequence 18, Appl
42	61	76.2	371	22	US-09-849-243-16	Sequence 16, Appl
43	60.5	75.6	102	8	US-08-467-551-2	Sequence 2, Appli
44	60.5	75.6	272	8	US-08-467-551-6	Sequence 6, Appli
45	60	75.0	2272	8	US-08-478-408-2	Sequence 2, Appli

ALIGNMENTS

```
RESULT   1
US-09-284-787-2
; Sequence 2, Application US/09284787
; GENERAL INFORMATION:
;  APPLICANT: Roche Diagnostics GmbH
;  TITLE OF INVENTION: Monoclonal Antibodies Against the Epitope YPYDVPDYA, A Process
;  TITLE OF INVENTION:  Producing the Same and their Use
;  FILE REFERENCE: BMID9913US
;  CURRENT APPLICATION NUMBER: US/09/284,787
;  CURRENT FILING DATE:  1999-08-16
;  PRIOR APPLICATION NUMBER: PCT/EP97/05783
;  PRIOR FILING DATE: 1997-10-20
;  NUMBER OF SEQ ID NOS: 3
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
;   LENGTH: 14
;   TYPE: PRT
;   ORGANISM: human influenza virus
US-09-284-787-2


  Query Match             100.0%;  Score 80;  DB 16;  Length 14;
  Best Local Similarity   100.0%;  Pred. No. 4.1e-06;
  Matches   14;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 YPYDVPDYAGSGSK 14
           ||||||||||||||
Db       1 YPYDVPDYAGSGSK 14


RESULT   2
US-09-780-224A-11
```

```
; Sequence 11, Application US/09780224A
; GENERAL INFORMATION:
;  APPLICANT: Roelvink, Petrus W
;  APPLICANT:  Kovesdi, Imre
;  APPLICANT:  Wickham, Thomas J
;  TITLE OF INVENTION: ADENOVIRAL CAPSID CONTAINING CHIMERIC PROTEIN IX
;  FILE REFERENCE: 208859
;  CURRENT APPLICATION NUMBER: US/09/780,224A
;  CURRENT FILING DATE:  2001-02-09
;  PRIOR APPLICATION NUMBER: US 60/181,163
;  PRIOR FILING DATE: 2000-02-09
;  NUMBER OF SEQ ID NOS: 13
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 11
;   LENGTH: 165
;   TYPE: PRT
;   ORGANISM: Adenovirus
US-09-780-224A-11


  Query Match             93.8%;  Score 75;  DB 21;  Length 165;
  Best Local Similarity   100.0%;  Pred. No. 0.00047;
  Matches   13;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 YPYDVPDYAGSGS 13
           |||||||||||||
Db       4 YPYDVPDYAGSGS 16


RESULT   3
US-09-475-158-56
; Sequence 56, Application US/09475158
; GENERAL INFORMATION:
;  APPLICANT: Gardella, Thomas J.
;  APPLICANT:  Kronenberg, Henry M.
;  APPLICANT:  Potts, John T.
;  APPLICANT:  Juppner, Harald
;  TITLE OF INVENTION: PTH Functional Domain Conjugate Peptides, Derivatives
;  TITLE OF INVENTION:  Thereof and Novel Tethered Ligand-Receptor Molecules
;  FILE REFERENCE: 0609.4780001
;  CURRENT APPLICATION NUMBER: US/09/475,158
;  CURRENT FILING DATE:  1999-12-30
;  PRIOR APPLICATION NUMBER: US 60/114,577
;  PRIOR FILING DATE: 1998-12-31
;  NUMBER OF SEQ ID NOS: 67
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 56
;   LENGTH: 32
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: modified PTH
;   OTHER INFORMATION: sequence
US-09-475-158-56


  Query Match             85.0%;  Score 68;  DB 18;  Length 32;
  Best Local Similarity   78.6%;  Pred. No. 0.001;
  Matches   11;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 YPYDVPDYAGSGSK 14
           |||||||||| | :
Db      10 YPYDVPDYAGGGGE 23


RESULT   4
US-09-622-646-23
; Sequence 23, Application US/09622646
; GENERAL INFORMATION:
;  APPLICANT: OZAKI, YASUKO
;  APPLICANT:  KOISHIHARA, YASUO
;  TITLE OF INVENTION: IMMUNOCHEMICAL ASSAY FOR ANTI-HM1.24 ANTIBODY
;  FILE REFERENCE: 053466/0286
;  CURRENT APPLICATION NUMBER: US/09/622,646
;  CURRENT FILING DATE:  2001-05-30
;  PRIOR APPLICATION NUMBER: PCT/JP99/00885
;  PRIOR FILING DATE: 1999-02-25
;  PRIOR APPLICATION NUMBER: JP 10-60613
;  PRIOR FILING DATE: 1998-02-25
;  NUMBER OF SEQ ID NOS: 29
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 23
;   LENGTH: 126
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: Amino acid sequence
;   OTHER INFORMATION: of a fusion protein comprising HA peptide and
;   OTHER INFORMATION: C-terminal-lacking soluble HM 1.24 antigenic protein
US-09-622-646-23


  Query Match             83.8%;  Score 67;  DB 20;  Length 126;
  Best Local Similarity   78.6%;  Pred. No. 0.0073;
  Matches   11;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 YPYDVPDYAGSGSK 14
           ||||||||||: |:
Db       1 YPYDVPDYAGTNSE 14


RESULT   5
US-09-622-646-22
; Sequence 22, Application US/09622646
; GENERAL INFORMATION:
;  APPLICANT: OZAKI, YASUKO
;  APPLICANT:  KOISHIHARA, YASUO
;  TITLE OF INVENTION: IMMUNOCHEMICAL ASSAY FOR ANTI-HM1.24 ANTIBODY
;  FILE REFERENCE: 053466/0286
;  CURRENT APPLICATION NUMBER: US/09/622,646
;  CURRENT FILING DATE:  2001-05-30
;  PRIOR APPLICATION NUMBER: PCT/JP99/00885
;  PRIOR FILING DATE: 1999-02-25
;  PRIOR APPLICATION NUMBER: JP 10-60613
;  PRIOR FILING DATE: 1998-02-25
;  NUMBER OF SEQ ID NOS: 29
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 22
;   LENGTH: 143
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: Amino acid sequence
;   OTHER INFORMATION: of a fusion protein comprising HA peptide and soluble
;   OTHER INFORMATION: HM 1.24 antigenic protein
US-09-622-646-22


  Query Match             83.8%;  Score 67;  DB 20;  Length 143;
  Best Local Similarity   78.6%;  Pred. No. 0.0084;
  Matches   11;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 YPYDVPDYAGSGSK 14
           ||||||||||: |:
Db       1 YPYDVPDYAGTNSE 14


RESULT   6
US-09-622-646-25
; Sequence 25, Application US/09622646
; GENERAL INFORMATION:
;  APPLICANT: OZAKI, YASUKO
;  APPLICANT:  KOISHIHARA, YASUO
;  TITLE OF INVENTION: IMMUNOCHEMICAL ASSAY FOR ANTI-HM1.24 ANTIBODY
;  FILE REFERENCE: 053466/0286
;  CURRENT APPLICATION NUMBER: US/09/622,646
```

```
;  CURRENT FILING DATE:  2001-05-30
;  PRIOR APPLICATION NUMBER: PCT/JP99/00885
;  PRIOR FILING DATE: 1999-02-25
;  PRIOR APPLICATION NUMBER: JP 10-60613
;  PRIOR FILING DATE: 1998-02-25
;  NUMBER OF SEQ ID NOS: 29
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 25
;   LENGTH: 147
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: Amino acid sequence
;   OTHER INFORMATION: of CG M/HA-HM164
US-09-622-646-25


  Query Match             83.8%;  Score 67;  DB 20;  Length 147;
  Best Local Similarity   78.6%;  Pred. No. 0.0087;
  Matches   11;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 YPYDVPDYAGSGSK 14
           |||||||||| : |:
Db      22 YPYDVPDYAGTNSE 35


RESULT   7
PCT-US00-09002-11
; Sequence 11, Application PC/TUS0009002
; GENERAL INFORMATION:
;  APPLICANT: Mayo Foundation for Medical Education and Research et al.
;  TITLE OF INVENTION: Anti-apoptotic fusion polypeptide
;  FILE REFERENCE: 150.195WO1
;  CURRENT APPLICATION NUMBER: PCT/US00/09002
;  CURRENT FILING DATE: 2000-04-07
;  EARLIER APPLICATION NUMBER: US 60/128,021
;  EARLIER FILING DATE: 1999-04-06
;  EARLIER APPLICATION NUMBER: US 60/127,867
;  EARLIER FILING DATE: 1999-04-05
;  NUMBER OF SEQ ID NOS: 13
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 11
;   LENGTH: 16
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: pTAT-HA linker
PCT-US00-09002-11


  Query Match             81.2%;  Score 65;  DB 1;  Length 16;
  Best Local Similarity   100.0%;  Pred. No. 0.0014;
  Matches   11;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 YPYDVPDYAGS 11
           |||||||||||
Db       4 YPYDVPDYAGS 14


RESULT   8
PCT-US98-16887-7
; Sequence 7, Application PC/TUS9816887
;  GENERAL INFORMATION:
;     APPLICANT:  Washington University
;     TITLE OF INVENTION:  INDUCIBLE REGULATORY SYSTEM
;     TITLE OF INVENTION:  AND USE THEREOF
;     NUMBER OF SEQUENCES:  8
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  Dike, Bronstein, Roberts & Cushman, LLP
;       STREET:  130 Water Street
;       CITY:  Boston
;       STATE:  MA
;       COUNTRY:  USA
;       ZIP:  02109
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Diskette
;       COMPUTER:  IBM Compatible
;       OPERATING SYSTEM:  DOS
;       SOFTWARE:  FastSEQ for Windows Version 2.0
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  PCT/US98/16887
;       FILING DATE:
;       CLASSIFICATION:
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  60/056,713
;       FILING DATE:  22-AUG-1997
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Corless, Peter F
;       REGISTRATION NUMBER:  33,860
;       REFERENCE/DOCKET NUMBER:  47275-PCT
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  617-523-3400
;       TELEFAX:  617-523-6440
;       TELEX:
;  INFORMATION FOR SEQ ID NO:  7:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  16 amino acids
;       TYPE:  amino acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  protein
PCT-US98-16887-7


  Query Match             81.2%;  Score 65;  DB 1;  Length 16;
  Best Local Similarity   100.0%;  Pred. No. 0.0014;
  Matches   11;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 YPYDVPDYAGS 11
           |||||||||||
Db       4 YPYDVPDYAGS 14


RESULT   9
US-09-134-793-7
; Sequence 7, Application US/09134793
;  GENERAL INFORMATION:
;     APPLICANT:  Dowdy, Steven F.
;     APPLICANT:  Jessee, Joel A.
;     TITLE OF INVENTION:  INDUCIBLE REGULATORY SYSTEM
;     TITLE OF INVENTION:  AND USE THEREOF
;     NUMBER OF SEQUENCES:  8
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  Dike, Bronstein, Roberts & Cushman, LLP
;       STREET:  130 Water Street
;       CITY:  Boston
;       STATE:  MA
;       COUNTRY:  USA
;       ZIP:  02109
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Diskette
;       COMPUTER:  IBM Compatible
;       OPERATING SYSTEM:  DOS
;       SOFTWARE:  FastSEQ for Windows Version 2.0
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/09/134,793
;       FILING DATE:
;       CLASSIFICATION:
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  60/056,713
;       FILING DATE:  22-AUG-1997
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Corless, Peter F
;       REGISTRATION NUMBER:  33,860
;       REFERENCE/DOCKET NUMBER:  47275
;     TELECOMMUNICATION INFORMATION:
```

```
;      TELEPHONE:  617-523-3400
;      TELEFAX:  617-523-6440
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  7:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  16 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-09-134-793-7


  Query Match             81.2%;  Score 65;  DB 15;  Length 16;
  Best Local Similarity   100.0%;  Pred. No. 0.0014;
  Matches   11;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 YPYDVPDYAGS 11
          ||||||||||
Db      4 YPYDVPDYAGS 14


RESULT  10
US-09-134-793A-7
; Sequence 7, Application US/09134793A
; GENERAL INFORMATION:
;  APPLICANT: DOWDY, STEVEN F.
;  APPLICANT:  JESSEE, JOEL A.
;  TITLE OF INVENTION: INDUCIBLE REGULATORY SYSTEM AND USE THEREOF
;  FILE REFERENCE: 1742/47275
;  CURRENT APPLICATION NUMBER: US/09/134,793A
;  CURRENT FILING DATE:  1998-08-14
;  NUMBER OF SEQ ID NOS: 9
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
;   LENGTH: 16
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: pTAT-HA linker
US-09-134-793A-7


  Query Match             81.2%;  Score 65;  DB 15;  Length 16;
  Best Local Similarity   100.0%;  Pred. No. 0.0014;
  Matches   11;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 YPYDVPDYAGS 11
          ||||||||||
Db      4 YPYDVPDYAGS 14


RESULT  11
US-09-300-639B-4
; Sequence 4, Application US/09300639B
; GENERAL INFORMATION:
;  APPLICANT:  Steven F. Dowdy
;  TITLE OF INVENTION:  Methods for Transducing Fusion Molecules
;  FILE REFERENCE:  48031
;  CURRENT APPLICATION NUMBER: US/09/300,639B
;  CURRENT FILING DATE:  1999-04-27
;  PRIOR APPLICATION NUMBER:  60/083,380
;  PRIOR FILING DATE:  1998-04-28
;  NUMBER OF SEQ ID NOS:  8
;  SOFTWARE:  FastSEQ for Windows Version 3.0
; SEQ ID NO 4
;   LENGTH: 16
;   TYPE: PRT
;   ORGANISM: Unknown
;   FEATURE:
;   OTHER INFORMATION: Amino acid sequence of pTAT-HA plasmid linker sequence
US-09-300-639B-4


  Query Match             81.2%;  Score 65;  DB 17;  Length 16;
  Best Local Similarity   100.0%;  Pred. No. 0.0014;
  Matches   11;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 YPYDVPDYAGS 11
          ||||||||||
Db      4 YPYDVPDYAGS 14


RESULT  12
US-09-459-113A-15
; Sequence 15, Application US/09459113A
; GENERAL INFORMATION:
;  APPLICANT:  Steven F. Dowdy
;  TITLE OF INVENTION:  PROTEIN TRANSDUCTION SYSTEM AND METHODS
;  TITLE OF INVENTION:  OF USE THEREOF
;  FILE REFERENCE:  48,884 (1742)
;  CURRENT APPLICATION NUMBER: US/09/459,113A
;  CURRENT FILING DATE:  1999-12-10
;  PRIOR APPLICATION NUMBER:  60/111,701
;  PRIOR FILING DATE:  1998-12-10
;  NUMBER OF SEQ ID NOS:  56
;  SOFTWARE:  FastSEQ for Windows Version 3.0
; SEQ ID NO 15
;   LENGTH: 16
;   TYPE: PRT
;   ORGANISM: Unknown
;   FEATURE:
;   OTHER INFORMATION: Synthetic pTAT-HA linker
US-09-459-113A-15


  Query Match             81.2%;  Score 65;  DB 18;  Length 16;
  Best Local Similarity   100.0%;  Pred. No. 0.0014;
  Matches   11;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 YPYDVPDYAGS 11
          ||||||||||
Db      4 YPYDVPDYAGS 14


RESULT  13
US-09-514-673-15
; Sequence 15, Application US/09514673
; GENERAL INFORMATION:
;  APPLICANT: DOWDY, STEVEN F.
;  TITLE OF INVENTION: NOVEL TRANSDUCTION MOLECULES AND METHODS FOR USING SAME
;  FILE REFERENCE: 49054(71742)
;  CURRENT APPLICATION NUMBER: US/09/514,673
;  CURRENT FILING DATE:  2000-02-28
;  PRIOR APPLICATION NUMBER: 60/122,757
;  PRIOR FILING DATE: 1999-02-28
;  PRIOR APPLICATION NUMBER: 60/151,291
;  PRIOR FILING DATE: 1999-08-29
;  NUMBER OF SEQ ID NOS: 59
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 15
;   LENGTH: 16
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: pTAT-HA linker
US-09-514-673-15


  Query Match             81.2%;  Score 65;  DB 19;  Length 16;
  Best Local Similarity   100.0%;  Pred. No. 0.0014;
  Matches   11;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 YPYDVPDYAGS 11
          ||||||||||
Db      4 YPYDVPDYAGS 14
```

```
RESULT  14
US-08-787-738-51
; Sequence 51, Application US/08787738
; GENERAL INFORMATION:
;  APPLICANT: Nolan, Garry P
;  APPLICANT:  Rothenburg, Michael S.
;  TITLE OF INVENTION: METHODS FOR SCREENING FOR TRANSDOMINANT EFFECTOR
;  TITLE OF INVENTION:  PEPTIDES AND RNA MOLECULES
;  FILE REFERENCE: A-64259-1 correction
;  CURRENT APPLICATION NUMBER: US/08/787,738
;  CURRENT FILING DATE:  2001-09-18
;  PRIOR APPLICATION NUMBER: 08/589,108
;  PRIOR FILING DATE: 1996-01-23
;  PRIOR APPLICATION NUMBER: 08/589,911
;  PRIOR FILING DATE: 1996-01-23
;  NUMBER OF SEQ ID NOS: 97
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 51
;   LENGTH: 17
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: synthetic.
US-08-787-738-51


  Query Match             81.2%;  Score 65;  DB 11;  Length 17;
  Best Local Similarity   100.0%;  Pred. No. 0.0015;
  Matches   11;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 YPYDVPDYAGS 11
          ||||||||||| |
Db      5 YPYDVPDYAGS 15


RESULT  15
US-08-787-738B-51
; Sequence 51, Application US/08787738B
; GENERAL INFORMATION:
;  APPLICANT: Nolan, Garry P
;  APPLICANT:  Rothenburg, Michael S.
;  TITLE OF INVENTION: METHODS FOR SCREENING FOR TRANSDOMINANT EFFECTOR
;  TITLE OF INVENTION:  PEPTIDES AND RNA MOLECULES
;  FILE REFERENCE: A-64259-1 correction
;  CURRENT APPLICATION NUMBER: US/08/787,738B
;  CURRENT FILING DATE:  2001-09-18
;  PRIOR APPLICATION NUMBER: 08/589,108
;  PRIOR FILING DATE: 1996-01-23
;  PRIOR APPLICATION NUMBER: 08/589,911
;  PRIOR FILING DATE: 1996-01-23
;  NUMBER OF SEQ ID NOS: 97
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 51
;   LENGTH: 17
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: synthetic.
US-08-787-738B-51


  Query Match             81.2%;  Score 65;  DB 11;  Length 17;
  Best Local Similarity   100.0%;  Pred. No. 0.0015;
  Matches   11;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 YPYDVPDYAGS 11
          ||||||||||| |
Db      5 YPYDVPDYAGS 15
```

Search completed: April 11, 2002, 10:20:51
Job time: 401 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5

OM protein - protein search, using sw model

Run on: April 11, 2002, 10:21:36 ; Search time 39.37 Seconds
(without alignments)
26.171 Million cell updates/sec

Title: US-09-284-787-2
Perfect score: 80
Sequence: 1 YPYDVPDYAGSGSK 14

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 367557 seqs, 73597022 residues

Total number of hits satisfying chosen parameters: 367557

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending_Patents_AA_New:*
1: /cgn2_6/ptodata/2/paa/PCT_NEW_COMB.pep:*
2: /cgn2_6/ptodata/2/paa/US06_NEW_COMB.pep:*
3: /cgn2_6/ptodata/2/paa/US07_NEW_COMB.pep:*
4: /cgn2_6/ptodata/2/paa/US08_NEW_COMB.pep:*
5: /cgn2_6/ptodata/2/paa/US09_NEW_COMB.pep:*
6: /cgn2_6/ptodata/2/paa/US09_NEW_COMB.pep1:*
7: /cgn2_6/ptodata/2/paa/US10_NEW_COMB.pep:*
8: /cgn2_6/ptodata/2/paa/US60_NEW_COMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	71	88.8	34	6	US-09-896-915-8	Sequence 8, Appli
2	65	81.2	16	5	US-09-775-052A-15	Sequence 15, Appl
3	65	81.2	16	6	US-09-775-052-15	Sequence 15, Appl
4	65	81.2	17	6	US-09-916-940-51	Sequence 51, Appl
5	65	81.2	17	7	US-10-096-550-51	Sequence 51, Appl
6	63	78.8	86	5	US-09-892-208A-2	Sequence 2, Appli
7	63	78.8	86	7	US-10-028-247-2	Sequence 2, Appli
8	59	73.8	14	6	US-09-740-876-1	Sequence 1, Appli
9	59	73.8	19	5	US-09-415-765A-57	Sequence 57, Appl
10	59	73.8	19	5	US-09-626-580B-57	Sequence 57, Appl
11	59	73.8	19	5	US-09-626-581C-57	Sequence 57, Appl
12	59	73.8	26	6	US-09-481-620A-34	Sequence 34, Appl
13	59	73.8	27	6	US-09-620-955B-31	Sequence 31, Appl
14	59	73.8	27	7	US-10-002-244-33	Sequence 33, Appl
15	59	73.8	41	6	US-09-481-620A-24	Sequence 24, Appl
16	59	73.8	41	7	US-10-002-244-25	Sequence 25, Appl
17	59	73.8	389	5	US-09-832-161-14	Sequence 14, Appl
18	58.5	73.1	540	7	US-10-015-535-22	Sequence 22, Appl
19	58.5	73.1	540	7	US-10-015-535-30	Sequence 30, Appl
20	57	71.2	13	5	US-09-395-061-4	Sequence 4, Appli
21	55	68.8	9	1	PCT-US01-45076-9	Sequence 9, Appli
22	55	68.8	9	5	US-09-466-568-19	Sequence 19, Appl
23	55	68.8	9	5	US-09-931-381A-24	Sequence 24, Appl
24	55	68.8	9	5	US-09-936-969-9	Sequence 9, Appli
25	55	68.8	9	5	US-09-180-167A-17	Sequence 17, Appl
26	55	68.8	9	5	US-09-775-052A-17	Sequence 17, Appl
27	55	68.8	9	6	US-09-912-733-32	Sequence 32, Appl
28	55	68.8	9	6	US-09-574-735C-45	Sequence 45, Appl
29	55	68.8	9	6	US-09-852-370A-33	Sequence 33, Appl
30	55	68.8	9	6	US-09-906-393A-4	Sequence 4, Appli
31	55	68.8	9	6	US-09-775-052-17	Sequence 17, Appl
32	55	68.8	9	6	US-09-554-726A-53	Sequence 53, Appl
33	55	68.8	9	6	US-09-845-667-32	Sequence 32, Appl
34	55	68.8	9	6	US-09-995-847-4	Sequence 4, Appli
35	55	68.8	9	6	US-09-554-726A-53	Sequence 53, Appl
36	55	68.8	9	6	US-09-529-106A-16	Sequence 16, Appl
37	55	68.8	9	6	US-09-713-858-7	Sequence 7, Appli
38	55	68.8	9	6	US-09-714-471-7	Sequence 7, Appli
39	55	68.8	9	6	US-09-896-915-37	Sequence 37, Appl
40	55	68.8	9	6	US-09-598-347-7	Sequence 7, Appli
41	55	68.8	9	6	US-09-830-980-8	Sequence 8, Appli
42	55	68.8	9	6	US-09-848-813-5	Sequence 5, Appli
43	55	68.8	9	6	US-09-918-036-17	Sequence 17, Appl
44	55	68.8	9	6	US-09-922-226-40	Sequence 40, Appl
45	55	68.8	9	6	US-09-807-305A-3	Sequence 3, Appli

ALIGNMENTS

RESULT 1
US-09-896-915-8
; Sequence 8, Application US/09896915
; GENERAL INFORMATION:
; APPLICANT: Medical Research Council
; APPLICANT: Griffiths, Andrew
; TITLE OF INVENTION: Optical Sorting Method
; FILE REFERENCE: 18396/2022
; CURRENT APPLICATION NUMBER: US/09/896,915
; CURRENT FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: GB9900298.2
; PRIOR FILING DATE: 1999-01-07
; PRIOR APPLICATION NUMBER: PCT/GB00/00030
; PRIOR FILING DATE: 2000-01-06
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 8
; LENGTH: 34
; TYPE: PRT
; ORGANISM: plasmid pET-23d(FLAGHA)
US-09-896-915-8

Query Match 88.8%; Score 71; DB 6; Length 34;
Best Local Similarity 92.3%; Pred. No. 3.7e-05;
Matches 12; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy       1 YPYDVPDYAGSGS 13
           |||||||||| ||
Db      22 YPYDVPDYAGGGS 34
```

RESULT 2
US-09-775-052A-15
; Sequence 15, Application US/09775052A
; GENERAL INFORMATION:
; APPLICANT: Dowdy, Steven F.
; TITLE OF INVENTION: ANTI-PATHOGEN SYSTEM AND METHODS OF USE THEREOF
; FILE REFERENCE: 48881/1742
; CURRENT APPLICATION NUMBER: US/09/775,052A
; CURRENT FILING DATE: 2001-12-05
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US/09/208,966
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-12-10
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/069,012
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-12-10
; NUMBER OF SEQ ID NOS: 57
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 15

; LENGTH: 16
; TYPE: PRT
; ORGANISM: human
US-09-775-052A-15

Query Match 81.2%; Score 65; DB 5; Length 16;
Best Local Similarity 100.0%; Pred. No. 0.00016;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 YPYDVPDYAGS 11
||||||||||||
Db 4 YPYDVPDYAGS 14

RESULT 3
US-09-775-052-15
; Sequence 15, Application US/09775052
; GENERAL INFORMATION:
; APPLICANT: Dowdy, Steven F.
; TITLE OF INVENTION: ANTI-PATHOGEN SYSTEM AND METHODS OF USE THEREOF
; FILE REFERENCE: 48881/1742
; CURRENT APPLICATION NUMBER: US/09/775,052
; CURRENT FILING DATE: 2001-02-01
; PRIOR APPLICATION NUMBER: 09/208,966
; PRIOR FILING DATE: 1998-12-10
; PRIOR APPLICATION NUMBER: 60/082,402
; PRIOR FILING DATE: 1998-04-20
; PRIOR APPLICATION NUMBER: 60/069,012
; PRIOR FILING DATE: 1997-12-10
; NUMBER OF SEQ ID NOS: 57
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 15
; LENGTH: 16
; TYPE: PRT
; ORGANISM: human
US-09-775-052-15

Query Match 81.2%; Score 65; DB 6; Length 16;
Best Local Similarity 100.0%; Pred. No. 0.00016;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 YPYDVPDYAGS 11
||||||||||||
Db 4 YPYDVPDYAGS 14

RESULT 4
US-09-916-940-51
; Sequence 51, Application US/09916940
; GENERAL INFORMATION:
; APPLICANT: Nolan, Garry P
; TITLE OF INVENTION: METHODS FOR SCREENING FOR TRANSDOMINANT INTRACELLULAR
; TITLE OF INVENTION: EFFECTOR PEPTIDES AND RNA MOLECULES
; FILE REFERENCE: A-64260-6/RMS/AMS
; CURRENT APPLICATION NUMBER: US/09/916,940
; CURRENT FILING DATE: 2001-07-27
; PRIOR APPLICATION NUMBER: US 09/727,715
; PRIOR FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: US 08/963,368
; PRIOR FILING DATE: 1997-11-03
; PRIOR APPLICATION NUMBER: US 08/589,109
; PRIOR FILING DATE: 1996-01-23
; PRIOR APPLICATION NUMBER: US 08/589,911
; PRIOR FILING DATE: 1996-01-23
; PRIOR APPLICATION NUMBER: US 08/789,333
; PRIOR FILING DATE: 1997-01-23
; PRIOR APPLICATION NUMBER: US 08/787,738
; PRIOR FILING DATE: 1997-01-23
; NUMBER OF SEQ ID NOS: 102
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 51
; LENGTH: 17
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: epitope tag
; OTHER INFORMATION: sequence.
US-09-916-940-51

Query Match 81.2%; Score 65; DB 6; Length 17;
Best Local Similarity 100.0%; Pred. No. 0.00017;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 YPYDVPDYAGS 11
||||||||||||
Db 5 YPYDVPDYAGS 15

RESULT 5
US-10-096-550-51
; Sequence 51, Application US/10096550
; GENERAL INFORMATION:
; APPLICANT: Nolan, Garry P
; APPLICANT: Rothenburg, Michael S.
; TITLE OF INVENTION: METHODS FOR SCREENING FOR TRANSDOMINANT EFFECTOR
; TITLE OF INVENTION: PEPTIDES AND RNA MOLECULES
; FILE REFERENCE: A-64259-1 correction
; CURRENT APPLICATION NUMBER: US/10/096,550
; CURRENT FILING DATE: 2002-03-12
; PRIOR APPLICATION NUMBER: 08/787,738
; PRIOR FILING DATE: 2001-09-18
; PRIOR APPLICATION NUMBER: 08/589,108
; PRIOR FILING DATE: 1996-01-23
; PRIOR APPLICATION NUMBER: 08/589,911
; PRIOR FILING DATE: 1996-01-23
; NUMBER OF SEQ ID NOS: 97
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 51
; LENGTH: 17
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic.
US-10-096-550-51

Query Match 81.2%; Score 65; DB 7; Length 17;
Best Local Similarity 100.0%; Pred. No. 0.00017;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 YPYDVPDYAGS 11
||||||||||||
Db 5 YPYDVPDYAGS 15

RESULT 6
US-09-892-208A-2
; Sequence 2, Application US/09892208A
; GENERAL INFORMATION:
; APPLICANT: Goldman, Stanley
; APPLICANT: Lathrop, Stephanie J.
; APPLICANT: Longchamp, Pascal F.
; APPLICANT: Whalen, Robert G.
; APPLICANT: Maxygen, Inc.
; TITLE OF INVENTION: Methods and Compositions for Developing Spore Display
; TITLE OF INVENTION: Systems for Medicinal and Industrial Applications
; FILE REFERENCE: 18097A-033510US
; CURRENT APPLICATION NUMBER: US/09/892,208A
; CURRENT FILING DATE: 2001-06-26
; PRIOR APPLICATION NUMBER: US 60/214,161
; PRIOR FILING DATE: 2000-06-26
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1

```
; SEQ ID NO 2
;   LENGTH: 86
;   TYPE: PRT
;   ORGANISM: Bacillus subtilis
;   FEATURE:
;   OTHER INFORMATION: CotC27 including HA11 epitope region
US-09-892-208A-2


  Query Match             78.8%;  Score 63;  DB 5;  Length 86;
  Best Local Similarity   84.6%;  Pred. No. 0.0021;
  Matches   11;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 YPYDVPDYAGSGS 13
          |||||||||  ||
Db     33 YPYDVPDYASLGS 45


RESULT   7
US-10-028-247-2
; Sequence 2, Application US/10028247
; GENERAL INFORMATION:
;  APPLICANT: Goldman, Stanley
;  APPLICANT:  Lathrop, Stephanie J.
;  APPLICANT:  Longchamp, Pascal F.
;  APPLICANT:  Whalen, Robert G.
;  APPLICANT:  Maxygen, Inc.
;  TITLE OF INVENTION: Methods and Compositions for Developing Spore Display
;  TITLE OF INVENTION:  Systems for Medicinal and Industrial Applications
;  FILE REFERENCE: 18097A-033520US
;  CURRENT APPLICATION NUMBER: US/10/028,247
;  CURRENT FILING DATE:  2001-12-19
;  PRIOR APPLICATION NUMBER: US 60/214,161
;  PRIOR FILING DATE: 2000-06-26
;  PRIOR APPLICATION NUMBER: US 09/892,208
;  PRIOR FILING DATE: 2001-06-26
;  NUMBER OF SEQ ID NOS: 6
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
;   LENGTH: 86
;   TYPE: PRT
;   ORGANISM: Bacillus subtilis
;   FEATURE:
;   OTHER INFORMATION: CotC27 including HA11 epitope region
US-10-028-247-2


  Query Match             78.8%;  Score 63;  DB 7;  Length 86;
  Best Local Similarity   84.6%;  Pred. No. 0.0021;
  Matches   11;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 YPYDVPDYAGSGS 13
          |||||||||  ||
Db     33 YPYDVPDYASLGS 45


RESULT   8
US-09-740-876-1
; Sequence 1, Application US/09740876
;   GENERAL INFORMATION:
;        APPLICANT: BRANELLEC, Didier
;                   WALSH, Kenneth
;                   ISNER, Jeffrey M.
;                   DENEFLE, Patrice
;        TITLE OF INVENTION: VIRAL VECTORS AND THEIR USE FOR TREATING
;                            HYPERPROLIFERATIVE DISORDERS, IN PARTICULAR RESTENOSIS
;        NUMBER OF SEQUENCES: 3
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Rhone-Poulenc Rorer Inc.
;             STREET: Mail Drop 3C43, P.O. Box 5093
;             CITY: Collegeville
;             STATE: PA
;             COUNTRY: USA
;             ZIP: 19426-0997
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.30
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/740,876
;             FILING DATE: 21-Dec-2000
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: 08/723,726
;             FILING DATE: <Unknown>
;             APPLICATION NUMBER: FR 95-04234
;             FILING DATE: 31-MAR-1995
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Savitzky, Martin F.
;             REGISTRATION NUMBER: 29,699
;             REFERENCE/DOCKET NUMBER: ST95022A-US
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (610)454-3816
;             TELEFAX: (610)454-3808
;   INFORMATION FOR SEQ ID NO: 1:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 14 amino acids
;             TYPE: amino acid
;             STRANDEDNESS: <Unknown>
;             TOPOLOGY: linear
;        MOLECULE TYPE: peptide
;        FRAGMENT TYPE: internal
;        SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-740-876-1


  Query Match             73.8%;  Score 59;  DB 6;  Length 14;
  Best Local Similarity   83.3%;  Pred. No. 0.0014;
  Matches   10;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 YPYDVPDYAGSG 12
          |||||||||  |
Db      1 YPYDVPDYASLG 12


RESULT   9
US-09-415-765A-57
; Sequence 57, Application US/09415765A
; GENERAL INFORMATION:
;  APPLICANT: Anderson, David
;  TITLE OF INVENTION: Fusions of Scaffold Proteins with Random Peptide
;  TITLE OF INVENTION:  Libraries
;  FILE REFERENCE: A66900-1/RMS/SJR
;  CURRENT APPLICATION NUMBER: US/09/415,765A
;  CURRENT FILING DATE:  1999-10-08
;  PRIOR APPLICATION NUMBER: 09/169,015
;  PRIOR FILING DATE: 1998-10-08
;  NUMBER OF SEQ ID NOS: 63
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 57
;   LENGTH: 19
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:  synthetic
US-09-415-765A-57


  Query Match             73.8%;  Score 59;  DB 5;  Length 19;
  Best Local Similarity   83.3%;  Pred. No. 0.0019;
  Matches   10;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 YPYDVPDYAGSG 12
          |||||||||  |
Db      5 YPYDVPDYASLG 16
```

```
RESULT  10
US-09-626-580B-57
; Sequence 57, Application US/09626580B
; GENERAL INFORMATION:
;  APPLICANT: Anderson, David
;  TITLE OF INVENTION: Fusions of Scaffold Proteins with Random Peptide
;  TITLE OF INVENTION:  Libraries
;  FILE REFERENCE: A-66900-2/RMS
;  CURRENT APPLICATION NUMBER: US/09/626,580B
;  CURRENT FILING DATE:  2002-02-21
;  PRIOR APPLICATION NUMBER: 09/169,015
;  PRIOR FILING DATE: 1998-10-08
;  PRIOR APPLICATION NUMBER: 09/415,765
;  PRIOR FILING DATE: 1999-10-08
;  NUMBER OF SEQ ID NOS: 63
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 57
;   LENGTH: 19
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:  synthetic
US-09-626-580B-57


  Query Match             73.8%;  Score 59;  DB 5;  Length 19;
  Best Local Similarity   83.3%;  Pred. No. 0.0019;
  Matches   10;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 YPYDVPDYAGSG 12
          |||||||||  |
Db      5 YPYDVPDYASLG 16


RESULT  11
US-09-626-581C-57
; Sequence 57, Application US/09626581C
; GENERAL INFORMATION:
;  APPLICANT: Anderson, David
;  TITLE OF INVENTION: Fusions of Scaffold Proteins with Random Peptide
;  TITLE OF INVENTION:  Libraries
;  FILE REFERENCE: A-66900-3/RMS
;  CURRENT APPLICATION NUMBER: US/09/626,581C
;  CURRENT FILING DATE:  2000-07-27
;  PRIOR APPLICATION NUMBER: 09/169,015
;  PRIOR FILING DATE: 1998-10-08
;  PRIOR APPLICATION NUMBER: 09/415,765
;  PRIOR FILING DATE: 1999-10-08
;  NUMBER OF SEQ ID NOS: 63
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 57
;   LENGTH: 19
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:  synthetic
US-09-626-581C-57


  Query Match             73.8%;  Score 59;  DB 5;  Length 19;
  Best Local Similarity   83.3%;  Pred. No. 0.0019;
  Matches   10;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 YPYDVPDYAGSG 12
          |||||||||  |
Db      5 YPYDVPDYASLG 16


RESULT  12
US-09-481-620A-34
; Sequence 34, Application US/09481620A
; GENERAL INFORMATION:
;  APPLICANT: ARIAD Gene Therapeutics, Inc.
;  TITLE OF INVENTION: Rapamycin Based Regulation of Biological Events
;  FILE REFERENCE: 345B PCT
;  CURRENT APPLICATION NUMBER: US/09/481,620A
;  CURRENT FILING DATE:  2000-01-12
;  NUMBER OF SEQ ID NOS: 132
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 34
;   LENGTH: 26
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: 12CA5epitope--SV40NLS
;   NAME/KEY: PEPTIDE
;   LOCATION: (1)..(26)
US-09-481-620A-34


  Query Match             73.8%;  Score 59;  DB 6;  Length 26;
  Best Local Similarity   83.3%;  Pred. No. 0.0027;
  Matches   10;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 YPYDVPDYAGSG 12
          |||||||||  |
Db      5 YPYDVPDYASLG 16


RESULT  13
US-09-620-955B-31
; Sequence 31, Application US/09620955B
; GENERAL INFORMATION:
;  APPLICANT: Huston, James S.
;  APPLICANT:  Messer, Anne
;  APPLICANT:  Lecerf, Jean-Michel
;  TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR INHIBITING POLYPEPTIDE
;  TITLE OF INVENTION:  ACCUMULATION ASSOCIATED WITH NEUROLOGICAL DISORDERS
;  FILE REFERENCE: INR-004CP
;  CURRENT APPLICATION NUMBER: US/09/620,955B
;  CURRENT FILING DATE:  2000-07-21
;  PRIOR APPLICATION NUMBER: 60/146,047
;  PRIOR FILING DATE: 1999-07-27
;  NUMBER OF SEQ ID NOS: 45
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 31
;   LENGTH: 27
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: Synthetic
;   OTHER INFORMATION: construct
US-09-620-955B-31


  Query Match             73.8%;  Score 59;  DB 6;  Length 27;
  Best Local Similarity   71.4%;  Pred. No. 0.0028;
  Matches   10;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 YPYDVPDYAGSGSK 14
          |||||||||    |
Db      1 YPYDVPDYASKDGK 14


RESULT  14
US-10-002-244-33
; Sequence 33, Application US/10002244
; GENERAL INFORMATION:
;  APPLICANT: ARIAD Gene Therapeutics, Inc.
;  TITLE OF INVENTION: Use of Heterologous Transcription Factors in Gene Therapy
;  FILE REFERENCE: 346B USC1
;  CURRENT APPLICATION NUMBER: US/10/002,244
;  CURRENT FILING DATE:  2002-01-29
;  NUMBER OF SEQ ID NOS: 68
```

```
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 33
;   LENGTH: 27
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   NAME/KEY: PEPTIDE
;   LOCATION: (1)..(27)
;   OTHER INFORMATION: pSMTN3
US-10-002-244-33


  Query Match             73.8%;  Score 59;  DB 7;  Length 27;
  Best Local Similarity   83.3%;  Pred. No. 0.0028;
  Matches   10;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 YPYDVPDYAGSG 12
          |||||||||  |
Db      5 YPYDVPDYASLG 16


RESULT  15
US-09-481-620A-24
; Sequence 24, Application US/09481620A
; GENERAL INFORMATION:
;  APPLICANT: ARIAD Gene Therapeutics, Inc.
;  TITLE OF INVENTION: Rapamycin Based Regulation of Biological Events
;  FILE REFERENCE: 345B PCT
;  CURRENT APPLICATION NUMBER: US/09/481,620A
;  CURRENT FILING DATE:  2000-01-12
;  NUMBER OF SEQ ID NOS: 132
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 24
;   LENGTH: 41
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: 12CA5epitope--SV40NLS--ZFHD5'
;   NAME/KEY: PEPTIDE
;   LOCATION: (1)..(41)
US-09-481-620A-24


  Query Match             73.8%;  Score 59;  DB 6;  Length 41;
  Best Local Similarity   83.3%;  Pred. No. 0.0044;
  Matches   10;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 YPYDVPDYAGSG 12
          |||||||||  |
Db      5 YPYDVPDYASLG 16


Search completed: April 11, 2002, 10:21:36
Job time: 446 sec
```

GenCore version 4.5

OM protein - protein search, using sw model

Run on: April 11, 2002, 10:20:51 ; Search time 266.73 Seconds
(without alignments)
13.533 Million cell updates/sec

Title: US-09-284-787-3
Perfect score: 75
Sequence: 1 SGSGYPYDVPDYA 13

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 3148936 seqs, 277657034 residues

Total number of hits satisfying chosen parameters: 3148936

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending_Patents_AA_Main:*
1: /cgn2_6/ptodata/2/paa/PCTUS_COMB.pep:*
2: /cgn2_6/ptodata/2/paa/US06_COMB.pep:*
3: /cgn2_6/ptodata/2/paa/US07_COMB.pep:*
4: /cgn2_6/ptodata/2/paa/US080_COMB.pep:*
5: /cgn2_6/ptodata/2/paa/US081_COMB.pep:*
6: /cgn2_6/ptodata/2/paa/US082_COMB.pep:*
7: /cgn2_6/ptodata/2/paa/US083_COMB.pep:*
8: /cgn2_6/ptodata/2/paa/US084_COMB.pep:*
9: /cgn2_6/ptodata/2/paa/US085_COMB.pep:*
10: /cgn2_6/ptodata/2/paa/US086_COMB.pep:*
11: /cgn2_6/ptodata/2/paa/US087_COMB.pep:*
12: /cgn2_6/ptodata/2/paa/US088_COMB.pep:*
13: /cgn2_6/ptodata/2/paa/US089_COMB.pep:*
14: /cgn2_6/ptodata/2/paa/US090_COMB.pep:*
15: /cgn2_6/ptodata/2/paa/US091_COMB.pep:*
16: /cgn2_6/ptodata/2/paa/US092_COMB.pep:*
17: /cgn2_6/ptodata/2/paa/US093_COMB.pep:*
18: /cgn2_6/ptodata/2/paa/US094_COMB.pep:*
19: /cgn2_6/ptodata/2/paa/US095_COMB.pep:*
20: /cgn2_6/ptodata/2/paa/US096_COMB.pep:*
21: /cgn2_6/ptodata/2/paa/US097_COMB.pep:*
22: /cgn2_6/ptodata/2/paa/US098_COMB.pep:*
23: /cgn2_6/ptodata/2/paa/US099_COMB.pep:*
24: /cgn2_6/ptodata/2/paa/US60_COMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	75	100.0	13	16	US-09-284-787-3	Sequence 3, Appli
2	67	89.3	17	11	US-08-787-738-51	Sequence 51, Appl
3	67	89.3	17	11	US-08-787-738B-51	Sequence 51, Appl
4	67	89.3	17	13	US-08-963-368-51	Sequence 51, Appl
5	67	89.3	18	16	US-09-285-912-136	Sequence 136, App
6	67	89.3	18	16	US-09-285-912A-136	Sequence 136, App
7	67	89.3	37	16	US-09-285-912-130	Sequence 130, App
8	67	89.3	37	16	US-09-285-912A-130	Sequence 130, App
9	67	89.3	38	16	US-09-285-912-68	Sequence 68, Appl
10	67	89.3	38	16	US-09-285-912A-68	Sequence 68, Appl
11	65	86.7	16	1	PCT-US00-09002-11	Sequence 11, Appl
12	65	86.7	16	1	PCT-US98-16887-7	Sequence 7, Appli
13	65	86.7	16	15	US-09-134-793-7	Sequence 7, Appli
14	65	86.7	16	15	US-09-134-793A-7	Sequence 7, Appli
15	65	86.7	16	17	US-09-300-639B-4	Sequence 4, Appli
16	65	86.7	16	18	US-09-459-113A-15	Sequence 15, Appl
17	65	86.7	16	19	US-09-514-673-15	Sequence 15, Appl
18	64.5	86.0	165	21	US-09-780-224A-13	Sequence 13, Appl
19	62	82.7	28	5	US-08-108-351-4	Sequence 4, Appli
20	62	82.7	225	11	US-08-756-693-8	Sequence 8, Appli
21	62	82.7	225	11	US-08-756-693A-8	Sequence 8, Appli
22	62	82.7	346	16	US-09-215-569B-8	Sequence 8, Appli
23	61	81.3	10	22	US-09-849-243-3	Sequence 3, Appli
24	61	81.3	11	10	US-08-628-829-17	Sequence 17, Appl
25	61	81.3	11	22	US-09-849-243-2	Sequence 2, Appli
26	61	81.3	12	20	US-09-608-958-46	Sequence 46, Appl
27	61	81.3	12	22	US-09-849-243-1	Sequence 1, Appli
28	61	81.3	21	7	US-08-361-811-127	Sequence 127, App
29	61	81.3	21	7	US-08-381-478-169	Sequence 169, App
30	61	81.3	21	9	US-08-535-025-169	Sequence 169, App
31	61	81.3	21	9	US-08-591-919-169	Sequence 169, App
32	61	81.3	21	9	US-08-592-007-169	Sequence 169, App
33	61	81.3	21	9	US-08-592-027-169	Sequence 169, App
34	61	81.3	21	9	US-08-592-131-169	Sequence 169, App
35	61	81.3	21	13	US-08-964-039-169	Sequence 169, App
36	61	81.3	80	18	US-09-457-568-18	Sequence 18, Appl
37	61	81.3	80	18	US-09-457-646-18	Sequence 18, Appl
38	61	81.3	80	19	US-09-516-065-18	Sequence 18, Appl
39	61	81.3	177	18	US-09-457-568-12	Sequence 12, Appl
40	61	81.3	177	18	US-09-457-646-12	Sequence 12, Appl
41	61	81.3	177	19	US-09-516-065-12	Sequence 12, Appl
42	61	81.3	224	1	PCT-US00-27237-36	Sequence 36, Appl
43	61	81.3	224	1	PCT-US00-27237-40	Sequence 40, Appl
44	61	81.3	224	1	PCT-US00-27237-44	Sequence 44, Appl
45	61	81.3	224	1	PCT-US00-27237-46	Sequence 46, Appl

ALIGNMENTS

RESULT 1
US-09-284-787-3
; Sequence 3, Application US/09284787
; GENERAL INFORMATION:
; APPLICANT: Roche Diagnostics GmbH
; TITLE OF INVENTION: Monoclonal Antibodies Against the Epitope YPYDVPDYA, A Process
; TITLE OF INVENTION: Producing the Same and their Use
; FILE REFERENCE: BMID9913US
; CURRENT APPLICATION NUMBER: US/09/284,787
; CURRENT FILING DATE: 1999-08-16
; PRIOR APPLICATION NUMBER: PCT/EP97/05783
; PRIOR FILING DATE: 1997-10-20
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 3
; LENGTH: 13
; TYPE: PRT
; ORGANISM: human influenza virus
US-09-284-787-3

Query Match 100.0%; Score 75; DB 16; Length 13;
Best Local Similarity 100.0%; Pred. No. 2.4e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       1 SGSGYPYDVPDYA 13
           |||||||||||||
Db       1 SGSGYPYDVPDYA 13
```

RESULT 2
US-08-787-738-51

```
; Sequence 51, Application US/08787738
; GENERAL INFORMATION:
;  APPLICANT: Nolan, Garry P
;  APPLICANT:  Rothenburg, Michael S.
;  TITLE OF INVENTION: METHODS FOR SCREENING FOR TRANSDOMINANT EFFECTOR
;  TITLE OF INVENTION:  PEPTIDES AND RNA MOLECULES
;  FILE REFERENCE: A-64259-1 correction
;  CURRENT APPLICATION NUMBER: US/08/787,738
;  CURRENT FILING DATE:  2001-09-18
;  PRIOR APPLICATION NUMBER: 08/589,108
;  PRIOR FILING DATE: 1996-01-23
;  PRIOR APPLICATION NUMBER: 08/589,911
;  PRIOR FILING DATE: 1996-01-23
;  NUMBER OF SEQ ID NOS: 97
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 51
;   LENGTH: 17
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: synthetic.
US-08-787-738-51


  Query Match             89.3%;  Score 67;  DB 11;  Length 17;
  Best Local Similarity   91.7%;  Pred. No. 0.00068;
  Matches   11;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      2 GSGYPYDVPDYA 13
          | ||||||||||
Db      2 GGGYPYDVPDYA 13


RESULT   3
US-08-787-738B-51
; Sequence 51, Application US/08787738B
; GENERAL INFORMATION:
;  APPLICANT: Nolan, Garry P
;  APPLICANT:  Rothenburg, Michael S.
;  TITLE OF INVENTION: METHODS FOR SCREENING FOR TRANSDOMINANT EFFECTOR
;  TITLE OF INVENTION:  PEPTIDES AND RNA MOLECULES
;  FILE REFERENCE: A-64259-1 correction
;  CURRENT APPLICATION NUMBER: US/08/787,738B
;  CURRENT FILING DATE:  2001-09-18
;  PRIOR APPLICATION NUMBER: 08/589,108
;  PRIOR FILING DATE: 1996-01-23
;  PRIOR APPLICATION NUMBER: 08/589,911
;  PRIOR FILING DATE: 1996-01-23
;  NUMBER OF SEQ ID NOS: 97
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 51
;   LENGTH: 17
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: synthetic.
US-08-787-738B-51


  Query Match             89.3%;  Score 67;  DB 11;  Length 17;
  Best Local Similarity   91.7%;  Pred. No. 0.00068;
  Matches   11;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      2 GSGYPYDVPDYA 13
          | ||||||||||
Db      2 GGGYPYDVPDYA 13


RESULT   4
US-08-963-368-51
; Sequence 51, Application US/08963368
; GENERAL INFORMATION:
;  APPLICANT:  Nolan, Garry P
;  TITLE OF INVENTION:  METHODS FOR SCREENING FOR TRANSDOMINANT INTRA-CELLULAR
;  TITLE OF INVENTION:  EFFECTOR PEPTIDES AND RNA MOLECULES
;  FILE REFERENCE:  A-64260-2/DJB/RMS/DR
;  CURRENT APPLICATION NUMBER: US/08/963,368
;  CURRENT FILING DATE:  1997-11-03
;  PRIOR APPLICATION NUMBER:  08/589,108
;  PRIOR FILING DATE:  1996-01-23
;  PRIOR APPLICATION NUMBER:  08/589,911
;  PRIOR FILING DATE:  1996-01-23
;  PRIOR APPLICATION NUMBER:  08/789,333
;  PRIOR FILING DATE:  1997-01-23
;  NUMBER OF SEQ ID NOS:  102
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 51
;   LENGTH: 17
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: epitope tag
;   OTHER INFORMATION: sequence.
US-08-963-368-51


  Query Match             89.3%;  Score 67;  DB 13;  Length 17;
  Best Local Similarity   91.7%;  Pred. No. 0.00068;
  Matches   11;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      2 GSGYPYDVPDYA 13
          | ||||||||||
Db      2 GGGYPYDVPDYA 13


RESULT   5
US-09-285-912-136
; Sequence 136, Application US/09285912
; GENERAL INFORMATION:
;  APPLICANT:  Anderson, David
;  TITLE OF INVENTION:  PEPTIDES CAUSING FORMATION OF COMPACT STRUCTURES
;  FILE REFERENCE:  A-66103-1/DJB/RMS/SJR
;  CURRENT APPLICATION NUMBER: US/09/285,912
;  CURRENT FILING DATE:  1999-04-02
;  PRIOR APPLICATION NUMBER:  60/080,444
;  PRIOR FILING DATE:  1998-04-02
;  NUMBER OF SEQ ID NOS:  150
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 136
;   LENGTH: 18
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-09-285-912-136


  Query Match             89.3%;  Score 67;  DB 16;  Length 18;
  Best Local Similarity   91.7%;  Pred. No. 0.00073;
  Matches   11;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      2 GSGYPYDVPDYA 13
          | ||||||||||
Db      2 GGGYPYDVPDYA 13


RESULT   6
US-09-285-912A-136
; Sequence 136, Application US/09285912A
; GENERAL INFORMATION:
;  APPLICANT: Anderson, David
;  TITLE OF INVENTION: PEPTIDES CAUSING FORMATION OF COMPACT STRUCTURES
;  FILE REFERENCE: A-66103-1/DJB/RMS/SJR
;  CURRENT APPLICATION NUMBER: US/09/285,912A
;  CURRENT FILING DATE:  1999-04-02
;  PRIOR APPLICATION NUMBER: 60/080,444
```

```
;  PRIOR FILING DATE: 1998-04-02
;  NUMBER OF SEQ ID NOS: 150
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 136
;   LENGTH: 18
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-09-285-912A-136


  Query Match             89.3%;  Score 67;  DB 16;  Length 18;
  Best Local Similarity   91.7%;  Pred. No. 0.00073;
  Matches   11;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      2 GSGYPYDVPDYA 13
          | ||||||||||
Db      2 GGGYPYDVPDYA 13


RESULT   7
US-09-285-912-130
; Sequence 130, Application US/09285912
; GENERAL INFORMATION:
;  APPLICANT:  Anderson, David
;  TITLE OF INVENTION:  PEPTIDES CAUSING FORMATION OF COMPACT STRUCTURES
;  FILE REFERENCE:  A-66103-1/DJB/RMS/SJR
;  CURRENT APPLICATION NUMBER: US/09/285,912
;  CURRENT FILING DATE:  1999-04-02
;  PRIOR APPLICATION NUMBER:  60/080,444
;  PRIOR FILING DATE:  1998-04-02
;  NUMBER OF SEQ ID NOS:  150
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 130
;   LENGTH: 37
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-09-285-912-130


  Query Match             89.3%;  Score 67;  DB 16;  Length 37;
  Best Local Similarity   91.7%;  Pred. No. 0.0017;
  Matches   11;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      2 GSGYPYDVPDYA 13
          | ||||||||||
Db     11 GGGYPYDVPDYA 22


RESULT   8
US-09-285-912A-130
; Sequence 130, Application US/09285912A
; GENERAL INFORMATION:
;  APPLICANT: Anderson, David
;  TITLE OF INVENTION: PEPTIDES CAUSING FORMATION OF COMPACT STRUCTURES
;  FILE REFERENCE: A-66103-1/DJB/RMS/SJR
;  CURRENT APPLICATION NUMBER: US/09/285,912A
;  CURRENT FILING DATE:  1999-04-02
;  PRIOR APPLICATION NUMBER: 60/080,444
;  PRIOR FILING DATE: 1998-04-02
;  NUMBER OF SEQ ID NOS: 150
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 130
;   LENGTH: 37
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-09-285-912A-130


  Query Match             89.3%;  Score 67;  DB 16;  Length 37;
  Best Local Similarity   91.7%;  Pred. No. 0.0017;
  Matches   11;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      2 GSGYPYDVPDYA 13
          | ||||||||||
Db     11 GGGYPYDVPDYA 22


RESULT   9
US-09-285-912-68
; Sequence 68, Application US/09285912
; GENERAL INFORMATION:
;  APPLICANT:  Anderson, David
;  TITLE OF INVENTION:  PEPTIDES CAUSING FORMATION OF COMPACT STRUCTURES
;  FILE REFERENCE:  A-66103-1/DJB/RMS/SJR
;  CURRENT APPLICATION NUMBER: US/09/285,912
;  CURRENT FILING DATE:  1999-04-02
;  PRIOR APPLICATION NUMBER:  60/080,444
;  PRIOR FILING DATE:  1998-04-02
;  NUMBER OF SEQ ID NOS:  150
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 68
;   LENGTH: 38
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-09-285-912-68


  Query Match             89.3%;  Score 67;  DB 16;  Length 38;
  Best Local Similarity   91.7%;  Pred. No. 0.0018;
  Matches   11;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      2 GSGYPYDVPDYA 13
          | ||||||||||
Db     11 GGGYPYDVPDYA 22


RESULT  10
US-09-285-912A-68
; Sequence 68, Application US/09285912A
; GENERAL INFORMATION:
;  APPLICANT: Anderson, David
;  TITLE OF INVENTION: PEPTIDES CAUSING FORMATION OF COMPACT STRUCTURES
;  FILE REFERENCE: A-66103-1/DJB/RMS/SJR
;  CURRENT APPLICATION NUMBER: US/09/285,912A
;  CURRENT FILING DATE:  1999-04-02
;  PRIOR APPLICATION NUMBER: 60/080,444
;  PRIOR FILING DATE: 1998-04-02
;  NUMBER OF SEQ ID NOS: 150
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 68
;   LENGTH: 38
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-09-285-912A-68


  Query Match             89.3%;  Score 67;  DB 16;  Length 38;
  Best Local Similarity   91.7%;  Pred. No. 0.0018;
  Matches   11;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      2 GSGYPYDVPDYA 13
          | ||||||||||
Db     11 GGGYPYDVPDYA 22


RESULT  11
```

```
PCT-US00-09002-11
; Sequence 11, Application PC/TUS0009002
; GENERAL INFORMATION:
;  APPLICANT: Mayo Foundation for Medical Education and Research et al.
;  TITLE OF INVENTION: Anti-apoptotic fusion polypeptide
;  FILE REFERENCE: 150.195WO1
;  CURRENT APPLICATION NUMBER: PCT/US00/09002
;  CURRENT FILING DATE: 2000-04-07
;  EARLIER APPLICATION NUMBER: US 60/128,021
;  EARLIER FILING DATE: 1999-04-06
;  EARLIER APPLICATION NUMBER: US 60/127,867
;  EARLIER FILING DATE: 1999-04-05
;  NUMBER OF SEQ ID NOS: 13
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 11
;   LENGTH: 16
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: pTAT-HA linker
PCT-US00-09002-11


  Query Match             86.7%;  Score 65;  DB 1;  Length 16;
  Best Local Similarity   100.0%;  Pred. No. 0.0014;
  Matches   11;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      3 SGYPYDVPDYA 13
          ||||||||||
Db      2 SGYPYDVPDYA 12


RESULT  12
PCT-US98-16887-7
; Sequence 7, Application PC/TUS9816887
;  GENERAL INFORMATION:
;    APPLICANT:  Washington University
;    TITLE OF INVENTION:  INDUCIBLE REGULATORY SYSTEM
;    TITLE OF INVENTION:  AND USE THEREOF
;    NUMBER OF SEQUENCES:  8
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Dike, Bronstein, Roberts & Cushman, LLP
;      STREET:  130 Water Street
;      CITY:  Boston
;      STATE:  MA
;      COUNTRY:  USA
;      ZIP:  02109
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US98/16887
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  60/056,713
;      FILING DATE:  22-AUG-1997
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Corless, Peter F
;      REGISTRATION NUMBER:  33,860
;      REFERENCE/DOCKET NUMBER:  47275-PCT
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617-523-3400
;      TELEFAX:  617-523-6440
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  7:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  16 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
PCT-US98-16887-7


  Query Match             86.7%;  Score 65;  DB 1;  Length 16;
  Best Local Similarity   100.0%;  Pred. No. 0.0014;
  Matches   11;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      3 SGYPYDVPDYA 13
          ||||||||||
Db      2 SGYPYDVPDYA 12


RESULT  13
US-09-134-793-7
; Sequence 7, Application US/09134793
;  GENERAL INFORMATION:
;    APPLICANT:  Dowdy, Steven F.
;    APPLICANT:  Jessee, Joel A.
;    TITLE OF INVENTION:  INDUCIBLE REGULATORY SYSTEM
;    TITLE OF INVENTION:  AND USE THEREOF
;    NUMBER OF SEQUENCES:  8
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Dike, Bronstein, Roberts & Cushman, LLP
;      STREET:  130 Water Street
;      CITY:  Boston
;      STATE:  MA
;      COUNTRY:  USA
;      ZIP:  02109
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/134,793
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  60/056,713
;      FILING DATE:  22-AUG-1997
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Corless, Peter F
;      REGISTRATION NUMBER:  33,860
;      REFERENCE/DOCKET NUMBER:  47275
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617-523-3400
;      TELEFAX:  617-523-6440
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  7:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  16 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-09-134-793-7


  Query Match             86.7%;  Score 65;  DB 15;  Length 16;
  Best Local Similarity   100.0%;  Pred. No. 0.0014;
  Matches   11;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      3 SGYPYDVPDYA 13
          ||||||||||
Db      2 SGYPYDVPDYA 12


RESULT  14
US-09-134-793A-7
; Sequence 7, Application US/09134793A
; GENERAL INFORMATION:
;  APPLICANT: DOWDY, STEVEN F.
```

```
;  APPLICANT:  JESSEE, JOEL A.
;  TITLE OF INVENTION: INDUCIBLE REGULATORY SYSTEM AND USE THEREOF
;  FILE REFERENCE: 1742/47275
;  CURRENT APPLICATION NUMBER: US/09/134,793A
;  CURRENT FILING DATE:  1998-08-14
;  NUMBER OF SEQ ID NOS: 9
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
;   LENGTH: 16
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: pTAT-HA linker
US-09-134-793A-7


  Query Match             86.7%;  Score 65;  DB 15;  Length 16;
  Best Local Similarity   100.0%;  Pred. No. 0.0014;
  Matches   11;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      3 SGYPYDVPDYA 13
          ||||||||||| 
Db      2 SGYPYDVPDYA 12


RESULT  15
US-09-300-639B-4
; Sequence 4, Application US/09300639B
; GENERAL INFORMATION:
;  APPLICANT:  Steven F. Dowdy
;  TITLE OF INVENTION:  Methods for Transducing Fusion Molecules
;  FILE REFERENCE:  48031
;  CURRENT APPLICATION NUMBER: US/09/300,639B
;  CURRENT FILING DATE:  1999-04-27
;  PRIOR APPLICATION NUMBER:  60/083,380
;  PRIOR FILING DATE:  1998-04-28
;  NUMBER OF SEQ ID NOS:  8
;  SOFTWARE:  FastSEQ for Windows Version 3.0
; SEQ ID NO 4
;   LENGTH: 16
;   TYPE: PRT
;   ORGANISM: Unknown
;   FEATURE:
;   OTHER INFORMATION: Amino acid sequence of pTAT-HA plasmid linker sequence
US-09-300-639B-4


  Query Match             86.7%;  Score 65;  DB 17;  Length 16;
  Best Local Similarity   100.0%;  Pred. No. 0.0014;
  Matches   11;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      3 SGYPYDVPDYA 13
          ||||||||||| 
Db      2 SGYPYDVPDYA 12


Search completed: April 11, 2002, 10:20:51
Job time: 401 sec
```

GenCore version 4.5

OM protein - protein search, using sw model

Run on: April 11, 2002, 10:21:36 ; Search time 39.37 Seconds
(without alignments)
24.302 Million cell updates/sec

Title: US-09-284-787-3
Perfect score: 75
Sequence: 1 SGSGYPYDVPDYA 13

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 367557 seqs, 73597022 residues

Total number of hits satisfying chosen parameters: 367557

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending_Patents_AA_New:*
1: /cgn2_6/ptodata/2/paa/PCT_NEW_COMB.pep:*
2: /cgn2_6/ptodata/2/paa/US06_NEW_COMB.pep:*
3: /cgn2_6/ptodata/2/paa/US07_NEW_COMB.pep:*
4: /cgn2_6/ptodata/2/paa/US08_NEW_COMB.pep:*
5: /cgn2_6/ptodata/2/paa/US09_NEW_COMB.pep:*
6: /cgn2_6/ptodata/2/paa/US09_NEW_COMB.pep1:*
7: /cgn2_6/ptodata/2/paa/US10_NEW_COMB.pep:*
8: /cgn2_6/ptodata/2/paa/US60_NEW_COMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	71	94.7	34	6	US-09-896-915-8	Sequence 8, Appli
2	67	89.3	17	6	US-09-916-940-51	Sequence 51, Appl
3	67	89.3	17	7	US-10-096-550-51	Sequence 51, Appl
4	67	89.3	19	5	US-09-415-765A-57	Sequence 57, Appl
5	67	89.3	19	5	US-09-626-580B-57	Sequence 57, Appl
6	67	89.3	19	5	US-09-626-581C-57	Sequence 57, Appl
7	65	86.7	16	5	US-09-775-052A-15	Sequence 15, Appl
8	65	86.7	16	6	US-09-775-052-15	Sequence 15, Appl
9	61	81.3	224	6	US-09-453-234-36	Sequence 36, Appl
10	61	81.3	224	6	US-09-453-234-40	Sequence 40, Appl
11	61	81.3	224	6	US-09-453-234-44	Sequence 44, Appl
12	61	81.3	224	6	US-09-453-234-46	Sequence 46, Appl
13	61	81.3	224	6	US-09-453-234-48	Sequence 48, Appl
14	61	81.3	224	6	US-09-453-234-52	Sequence 52, Appl
15	61	81.3	224	6	US-09-453-234-76	Sequence 76, Appl
16	61	81.3	224	6	US-09-453-234-78	Sequence 78, Appl
17	61	81.3	224	6	US-09-453-234-82	Sequence 82, Appl
18	61	81.3	224	6	US-09-453-234-84	Sequence 84, Appl
19	61	81.3	224	6	US-09-453-234-88	Sequence 88, Appl
20	61	81.3	224	6	US-09-453-234-90	Sequence 90, Appl
21	61	81.3	226	6	US-09-453-234-38	Sequence 38, Appl
22	61	81.3	226	6	US-09-453-234-42	Sequence 42, Appl
23	61	81.3	226	6	US-09-453-234-50	Sequence 50, Appl
24	61	81.3	226	6	US-09-453-234-72	Sequence 72, Appl
25	61	81.3	226	6	US-09-453-234-74	Sequence 74, Appl
26	61	81.3	226	6	US-09-453-234-80	Sequence 80, Appl
27	61	81.3	226	6	US-09-453-234-86	Sequence 86, Appl
28	59	78.7	26	6	US-09-481-620A-34	Sequence 34, Appl
29	59	78.7	27	7	US-10-002-244-33	Sequence 33, Appl
30	59	78.7	41	6	US-09-481-620A-24	Sequence 24, Appl
31	59	78.7	41	7	US-10-002-244-25	Sequence 25, Appl
32	59	78.7	86	5	US-09-892-208A-2	Sequence 2, Appli
33	59	78.7	86	7	US-10-028-247-2	Sequence 2, Appli
34	58	77.3	47	6	US-09-852-370A-65	Sequence 65, Appl
35	58	77.3	316	7	US-10-041-264A-9	Sequence 9, Appli
36	58	77.3	316	7	US-10-041-400A-9	Sequence 9, Appli
37	58	77.3	316	7	US-10-042-091A-9	Sequence 9, Appli
38	55	73.3	9	1	PCT-US01-45076-9	Sequence 9, Appli
39	55	73.3	9	5	US-09-466-568-19	Sequence 19, Appl
40	55	73.3	9	5	US-09-931-381A-24	Sequence 24, Appl
41	55	73.3	9	5	US-09-936-969-9	Sequence 9, Appli
42	55	73.3	9	5	US-09-180-167A-17	Sequence 17, Appl
43	55	73.3	9	5	US-09-775-052A-17	Sequence 17, Appl
44	55	73.3	9	6	US-09-912-733-32	Sequence 32, Appl
45	55	73.3	9	6	US-09-845-667-32	Sequence 32, Appl

ALIGNMENTS

RESULT 1
US-09-896-915-8
; Sequence 8, Application US/09896915
; GENERAL INFORMATION:
; APPLICANT: Medical Research Council
; APPLICANT: Griffiths, Andrew
; TITLE OF INVENTION: Optical Sorting Method
; FILE REFERENCE: 18396/2022
; CURRENT APPLICATION NUMBER: US/09/896,915
; CURRENT FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: GB9900298.2
; PRIOR FILING DATE: 1999-01-07
; PRIOR APPLICATION NUMBER: PCT/GB00/00030
; PRIOR FILING DATE: 2000-01-06
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 8
; LENGTH: 34
; TYPE: PRT
; ORGANISM: plasmid pET-23d(FLAGHA)
US-09-896-915-8

Query Match 94.7%; Score 71; DB 6; Length 34;
Best Local Similarity 92.3%; Pred. No. 4.2e-05;
Matches 12; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy       1 SGSGYPYDVPDYA 13
           || |||||||||||
Db      18 SGGGYPYDVPDYA 30
```

RESULT 2
US-09-916-940-51
; Sequence 51, Application US/09916940
; GENERAL INFORMATION:
; APPLICANT: Nolan, Garry P
; TITLE OF INVENTION: METHODS FOR SCREENING FOR TRANSDOMINANT INTRACELLULAR
; TITLE OF INVENTION: EFFECTOR PEPTIDES AND RNA MOLECULES
; FILE REFERENCE: A-64260-6/RMS/AMS
; CURRENT APPLICATION NUMBER: US/09/916,940
; CURRENT FILING DATE: 2001-07-27
; PRIOR APPLICATION NUMBER: US 09/727,715
; PRIOR FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: US 08/963,368
; PRIOR FILING DATE: 1997-11-03
; PRIOR APPLICATION NUMBER: US 08/589,109
; PRIOR FILING DATE: 1996-01-23

```
;  PRIOR APPLICATION NUMBER: US 08/589,911
;  PRIOR FILING DATE: 1996-01-23
;  PRIOR APPLICATION NUMBER: US 08/789,333
;  PRIOR FILING DATE: 1997-01-23
;  PRIOR APPLICATION NUMBER: US 08/787,738
;  PRIOR FILING DATE: 1997-01-23
;  NUMBER OF SEQ ID NOS: 102
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 51
;   LENGTH: 17
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: epitope tag
;   OTHER INFORMATION: sequence.
US-09-916-940-51


  Query Match             89.3%;  Score 67;  DB 6;  Length 17;
  Best Local Similarity   91.7%;  Pred. No. 9.2e-05;
  Matches   11;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      2 GSGYPYDVPDYA 13
          | ||||||||||
Db      2 GGGYPYDVPDYA 13


RESULT   3
US-10-096-550-51
; Sequence 51, Application US/10096550
; GENERAL INFORMATION:
;  APPLICANT: Nolan, Garry P
;  APPLICANT:  Rothenburg, Michael S.
;  TITLE OF INVENTION: METHODS FOR SCREENING FOR TRANSDOMINANT EFFECTOR
;  TITLE OF INVENTION:  PEPTIDES AND RNA MOLECULES
;  FILE REFERENCE: A-64259-1 correction
;  CURRENT APPLICATION NUMBER: US/10/096,550
;  CURRENT FILING DATE:  2002-03-12
;  PRIOR APPLICATION NUMBER: 08/787,738
;  PRIOR FILING DATE: 2001-09-18
;  PRIOR APPLICATION NUMBER: 08/589,108
;  PRIOR FILING DATE: 1996-01-23
;  PRIOR APPLICATION NUMBER: 08/589,911
;  PRIOR FILING DATE: 1996-01-23
;  NUMBER OF SEQ ID NOS: 97
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 51
;   LENGTH: 17
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: synthetic.
US-10-096-550-51


  Query Match             89.3%;  Score 67;  DB 7;  Length 17;
  Best Local Similarity   91.7%;  Pred. No. 9.2e-05;
  Matches   11;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      2 GSGYPYDVPDYA 13
          | ||||||||||
Db      2 GGGYPYDVPDYA 13


RESULT   4
US-09-415-765A-57
; Sequence 57, Application US/09415765A
; GENERAL INFORMATION:
;  APPLICANT: Anderson, David
;  TITLE OF INVENTION: Fusions of Scaffold Proteins with Random Peptide
;  TITLE OF INVENTION:  Libraries
;  FILE REFERENCE: A66900-1/RMS/SJR
;  CURRENT APPLICATION NUMBER: US/09/415,765A
;  CURRENT FILING DATE:  1999-10-08
;  PRIOR APPLICATION NUMBER: 09/169,015
;  PRIOR FILING DATE: 1998-10-08
;  NUMBER OF SEQ ID NOS: 63
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 57
;   LENGTH: 19
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:  synthetic
US-09-415-765A-57


  Query Match             89.3%;  Score 67;  DB 5;  Length 19;
  Best Local Similarity   91.7%;  Pred. No. 0.0001;
  Matches   11;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      2 GSGYPYDVPDYA 13
          | ||||||||||
Db      2 GGGYPYDVPDYA 13


RESULT   5
US-09-626-580B-57
; Sequence 57, Application US/09626580B
; GENERAL INFORMATION:
;  APPLICANT: Anderson, David
;  TITLE OF INVENTION: Fusions of Scaffold Proteins with Random Peptide
;  TITLE OF INVENTION:  Libraries
;  FILE REFERENCE: A-66900-2/RMS
;  CURRENT APPLICATION NUMBER: US/09/626,580B
;  CURRENT FILING DATE:  2002-02-21
;  PRIOR APPLICATION NUMBER: 09/169,015
;  PRIOR FILING DATE: 1998-10-08
;  PRIOR APPLICATION NUMBER: 09/415,765
;  PRIOR FILING DATE: 1999-10-08
;  NUMBER OF SEQ ID NOS: 63
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 57
;   LENGTH: 19
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:  synthetic
US-09-626-580B-57


  Query Match             89.3%;  Score 67;  DB 5;  Length 19;
  Best Local Similarity   91.7%;  Pred. No. 0.0001;
  Matches   11;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      2 GSGYPYDVPDYA 13
          | ||||||||||
Db      2 GGGYPYDVPDYA 13


RESULT   6
US-09-626-581C-57
; Sequence 57, Application US/09626581C
; GENERAL INFORMATION:
;  APPLICANT: Anderson, David
;  TITLE OF INVENTION: Fusions of Scaffold Proteins with Random Peptide
;  TITLE OF INVENTION:  Libraries
;  FILE REFERENCE: A-66900-3/RMS
;  CURRENT APPLICATION NUMBER: US/09/626,581C
;  CURRENT FILING DATE:  2000-07-27
;  PRIOR APPLICATION NUMBER: 09/169,015
;  PRIOR FILING DATE: 1998-10-08
;  PRIOR APPLICATION NUMBER: 09/415,765
;  PRIOR FILING DATE: 1999-10-08
;  NUMBER OF SEQ ID NOS: 63
;  SOFTWARE: PatentIn Ver. 2.0
```

; SEQ ID NO 57
; LENGTH: 19
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
US-09-626-581C-57

Query Match 89.3%; Score 67; DB 5; Length 19;
Best Local Similarity 91.7%; Pred. No. 0.0001;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 GSGYPYDVPDYA 13
| ||||||||||
Db 2 GGGYPYDVPDYA 13

RESULT 7
US-09-775-052A-15
; Sequence 15, Application US/09775052A
; GENERAL INFORMATION:
; APPLICANT: Dowdy, Steven F.
; TITLE OF INVENTION: ANTI-PATHOGEN SYSTEM AND METHODS OF USE THEREOF
; FILE REFERENCE: 48881/1742
; CURRENT APPLICATION NUMBER: US/09/775,052A
; CURRENT FILING DATE: 2001-12-05
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US/09/208,966
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-12-10
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/069,012
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-12-10
; NUMBER OF SEQ ID NOS: 57
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 15
; LENGTH: 16
; TYPE: PRT
; ORGANISM: human
US-09-775-052A-15

Query Match 86.7%; Score 65; DB 5; Length 16;
Best Local Similarity 100.0%; Pred. No. 0.00018;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 SGYPYDVPDYA 13
||||||||||||
Db 2 SGYPYDVPDYA 12

RESULT 8
US-09-775-052-15
; Sequence 15, Application US/09775052
; GENERAL INFORMATION:
; APPLICANT: Dowdy, Steven F.
; TITLE OF INVENTION: ANTI-PATHOGEN SYSTEM AND METHODS OF USE THEREOF
; FILE REFERENCE: 48881/1742
; CURRENT APPLICATION NUMBER: US/09/775,052
; CURRENT FILING DATE: 2001-02-01
; PRIOR APPLICATION NUMBER: 09/208,966
; PRIOR FILING DATE: 1998-12-10
; PRIOR APPLICATION NUMBER: 60/082,402
; PRIOR FILING DATE: 1998-04-20
; PRIOR APPLICATION NUMBER: 60/069,012
; PRIOR FILING DATE: 1997-12-10
; NUMBER OF SEQ ID NOS: 57
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 15
; LENGTH: 16
; TYPE: PRT
; ORGANISM: human
US-09-775-052-15

Query Match 86.7%; Score 65; DB 6; Length 16;
Best Local Similarity 100.0%; Pred. No. 0.00018;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 SGYPYDVPDYA 13
||||||||||||
Db 2 SGYPYDVPDYA 12

RESULT 9
US-09-453-234-36
; Sequence 36, Application US/09453234
; GENERAL INFORMATION:
; APPLICANT: Buechler, Joe
; APPLICANT: Valkirs, Gunars
; APPLICANT: Gray, Jeff
; APPLICANT: Lonberg, Nils
; APPLICANT: Biosite Diagnostics, Inc.
; APPLICANT: GenPharm International
; TITLE OF INVENTION: Human Antibodies
; FILE REFERENCE: 020015-000110US
; CURRENT APPLICATION NUMBER: US/09/453,234
; CURRENT FILING DATE: 1999-12-01
; PRIOR APPLICATION NUMBER: US 60/157,415
; PRIOR FILING DATE: 1999-10-02
; NUMBER OF SEQ ID NOS: 112
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 36
; LENGTH: 224
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-453-234-36

Query Match 81.3%; Score 61; DB 6; Length 224;
Best Local Similarity 83.3%; Pred. No. 0.013;
Matches 10; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 GSGYPYDVPDYA 13
| ||||||||||
Db 212 GESYPYDVPDYA 223

RESULT 10
US-09-453-234-40
; Sequence 40, Application US/09453234
; GENERAL INFORMATION:
; APPLICANT: Buechler, Joe
; APPLICANT: Valkirs, Gunars
; APPLICANT: Gray, Jeff
; APPLICANT: Lonberg, Nils
; APPLICANT: Biosite Diagnostics, Inc.
; APPLICANT: GenPharm International
; TITLE OF INVENTION: Human Antibodies
; FILE REFERENCE: 020015-000110US
; CURRENT APPLICATION NUMBER: US/09/453,234
; CURRENT FILING DATE: 1999-12-01
; PRIOR APPLICATION NUMBER: US 60/157,415
; PRIOR FILING DATE: 1999-10-02
; NUMBER OF SEQ ID NOS: 112
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 40
; LENGTH: 224
; TYPE: PRT
; ORGANISM: Homo sapiens
; OTHER INFORMATION: M1-4L
US-09-453-234-40

Query Match 81.3%; Score 61; DB 6; Length 224;
Best Local Similarity 83.3%; Pred. No. 0.013;
Matches 10; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy       2 GSGYPYDVPDYA 13
           |  |||||||||
Db     212 GESYPYDVPDYA 223
```

```
RESULT  11
US-09-453-234-44
; Sequence 44, Application US/09453234
; GENERAL INFORMATION:
;  APPLICANT: Buechler, Joe
;  APPLICANT:  Valkirs, Gunars
;  APPLICANT:  Gray, Jeff
;  APPLICANT:  Lonberg, Nils
;  APPLICANT:  Biosite Diagnostics, Inc.
;  APPLICANT:  GenPharm International
;  TITLE OF INVENTION: Human Antibodies
;  FILE REFERENCE: 020015-000110US
;  CURRENT APPLICATION NUMBER: US/09/453,234
;  CURRENT FILING DATE:  1999-12-01
;  PRIOR APPLICATION NUMBER: US 60/157,415
;  PRIOR FILING DATE: 1999-10-02
;  NUMBER OF SEQ ID NOS: 112
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 44
;   LENGTH: 224
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   OTHER INFORMATION: M1-8L
US-09-453-234-44
```

```
  Query Match            81.3%;  Score 61;  DB 6;  Length 224;
  Best Local Similarity  83.3%;  Pred. No. 0.013;
  Matches   10;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       2 GSGYPYDVPDYA 13
           |  |||||||||
Db     212 GESYPYDVPDYA 223
```

```
RESULT  12
US-09-453-234-46
; Sequence 46, Application US/09453234
; GENERAL INFORMATION:
;  APPLICANT: Buechler, Joe
;  APPLICANT:  Valkirs, Gunars
;  APPLICANT:  Gray, Jeff
;  APPLICANT:  Lonberg, Nils
;  APPLICANT:  Biosite Diagnostics, Inc.
;  APPLICANT:  GenPharm International
;  TITLE OF INVENTION: Human Antibodies
;  FILE REFERENCE: 020015-000110US
;  CURRENT APPLICATION NUMBER: US/09/453,234
;  CURRENT FILING DATE:  1999-12-01
;  PRIOR APPLICATION NUMBER: US 60/157,415
;  PRIOR FILING DATE: 1999-10-02
;  NUMBER OF SEQ ID NOS: 112
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 46
;   LENGTH: 224
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   OTHER INFORMATION: M1-10L
US-09-453-234-46
```

```
  Query Match            81.3%;  Score 61;  DB 6;  Length 224;
  Best Local Similarity  83.3%;  Pred. No. 0.013;
  Matches   10;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       2 GSGYPYDVPDYA 13
           |  |||||||||
Db     212 GESYPYDVPDYA 223
```

```
RESULT  13
US-09-453-234-48
; Sequence 48, Application US/09453234
; GENERAL INFORMATION:
;  APPLICANT: Buechler, Joe
;  APPLICANT:  Valkirs, Gunars
;  APPLICANT:  Gray, Jeff
;  APPLICANT:  Lonberg, Nils
;  APPLICANT:  Biosite Diagnostics, Inc.
;  APPLICANT:  GenPharm International
;  TITLE OF INVENTION: Human Antibodies
;  FILE REFERENCE: 020015-000110US
;  CURRENT APPLICATION NUMBER: US/09/453,234
;  CURRENT FILING DATE:  1999-12-01
;  PRIOR APPLICATION NUMBER: US 60/157,415
;  PRIOR FILING DATE: 1999-10-02
;  NUMBER OF SEQ ID NOS: 112
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 48
;   LENGTH: 224
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   OTHER INFORMATION: M1-21L
US-09-453-234-48
```

```
  Query Match            81.3%;  Score 61;  DB 6;  Length 224;
  Best Local Similarity  83.3%;  Pred. No. 0.013;
  Matches   10;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       2 GSGYPYDVPDYA 13
           |  |||||||||
Db     212 GESYPYDVPDYA 223
```

```
RESULT  14
US-09-453-234-52
; Sequence 52, Application US/09453234
; GENERAL INFORMATION:
;  APPLICANT: Buechler, Joe
;  APPLICANT:  Valkirs, Gunars
;  APPLICANT:  Gray, Jeff
;  APPLICANT:  Lonberg, Nils
;  APPLICANT:  Biosite Diagnostics, Inc.
;  APPLICANT:  GenPharm International
;  TITLE OF INVENTION: Human Antibodies
;  FILE REFERENCE: 020015-000110US
;  CURRENT APPLICATION NUMBER: US/09/453,234
;  CURRENT FILING DATE:  1999-12-01
;  PRIOR APPLICATION NUMBER: US 60/157,415
;  PRIOR FILING DATE: 1999-10-02
;  NUMBER OF SEQ ID NOS: 112
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 52
;   LENGTH: 224
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: M1-25L
US-09-453-234-52
```

```
  Query Match            81.3%;  Score 61;  DB 6;  Length 224;
  Best Local Similarity  83.3%;  Pred. No. 0.013;
  Matches   10;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       2 GSGYPYDVPDYA 13
           |  |||||||||
Db     212 GESYPYDVPDYA 223
```

```
RESULT  15
US-09-453-234-76
; Sequence 76, Application US/09453234
; GENERAL INFORMATION:
;  APPLICANT: Buechler, Joe
;  APPLICANT:  Valkirs, Gunars
;  APPLICANT:  Gray, Jeff
;  APPLICANT:  Lonberg, Nils
;  APPLICANT:  Biosite Diagnostics, Inc.
;  APPLICANT:  GenPharm International
;  TITLE OF INVENTION: Human Antibodies
;  FILE REFERENCE: 020015-000110US
;  CURRENT APPLICATION NUMBER: US/09/453,234
;  CURRENT FILING DATE:  1999-12-01
;  PRIOR APPLICATION NUMBER: US 60/157,415
;  PRIOR FILING DATE: 1999-10-02
;  NUMBER OF SEQ ID NOS: 112
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 76
;   LENGTH: 224
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   OTHER INFORMATION: M2-16L
US-09-453-234-76


  Query Match             81.3%;  Score 61;  DB 6;  Length 224;
  Best Local Similarity   83.3%;  Pred. No. 0.013;
  Matches   10;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       2 GSGYPYDVPDYA 13
           |  |||||||||
Db     212 GESYPYDVPDYA 223


Search completed: April 11, 2002, 10:21:36
Job time: 446 sec
```

GenCore version 4.5

OM protein - protein search, using sw model

Run on: April 11, 2002, 10:14:10 ; Search time 53.05 Seconds
(without alignments)
12.567 Million cell updates/sec

Title: US-09-284-787-1
Perfect score: 55
Sequence: 1 YPYDVPDYA 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 522463 seqs, 74073290 residues

Total number of hits satisfying chosen parameters: 522463

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_1101:*
1: /SIDS8/gcgdata/geneseq/geneseqp/AA1980.DAT:*
2: /SIDS8/gcgdata/geneseq/geneseqp/AA1981.DAT:*
3: /SIDS8/gcgdata/geneseq/geneseqp/AA1982.DAT:*
4: /SIDS8/gcgdata/geneseq/geneseqp/AA1983.DAT:*
5: /SIDS8/gcgdata/geneseq/geneseqp/AA1984.DAT:*
6: /SIDS8/gcgdata/geneseq/geneseqp/AA1985.DAT:*
7: /SIDS8/gcgdata/geneseq/geneseqp/AA1986.DAT:*
8: /SIDS8/gcgdata/geneseq/geneseqp/AA1987.DAT:*
9: /SIDS8/gcgdata/geneseq/geneseqp/AA1988.DAT:*
10: /SIDS8/gcgdata/geneseq/geneseqp/AA1989.DAT:*
11: /SIDS8/gcgdata/geneseq/geneseqp/AA1990.DAT:*
12: /SIDS8/gcgdata/geneseq/geneseqp/AA1991.DAT:*
13: /SIDS8/gcgdata/geneseq/geneseqp/AA1992.DAT:*
14: /SIDS8/gcgdata/geneseq/geneseqp/AA1993.DAT:*
15: /SIDS8/gcgdata/geneseq/geneseqp/AA1994.DAT:*
16: /SIDS8/gcgdata/geneseq/geneseqp/AA1995.DAT:*
17: /SIDS8/gcgdata/geneseq/geneseqp/AA1996.DAT:*
18: /SIDS8/gcgdata/geneseq/geneseqp/AA1997.DAT:*
19: /SIDS8/gcgdata/geneseq/geneseqp/AA1998.DAT:*
20: /SIDS8/gcgdata/geneseq/geneseqp/AA1999.DAT:*
21: /SIDS8/gcgdata/geneseq/geneseqp/AA2000.DAT:*
22: /SIDS8/gcgdata/geneseq/geneseqp/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	55	100.0	9	15	AAR56545	Hal epitope tag.
2	55	100.0	9	16	AAR78372	Viral antigenic ep
3	55	100.0	9	17	AAW03654	HA-epitope tag. S
4	55	100.0	9	17	AAR92029	Hydrophilic peptid
5	55	100.0	9	18	AAW30296	Epitope for anti-H
6	55	100.0	9	18	AAW27094	Transforming growt
7	55	100.0	9	18	AAW00970	Haemagglutinin epi
8	55	100.0	9	18	AAW23660	pJG4-5 library pla
9	55	100.0	9	18	AAW08162	Purification tag o
10	55	100.0	9	18	AAW12385	Influenza virus ha
11	55	100.0	9	18	AAW01661	10-mer peptide tag
12	55	100.0	9	19	AAW81556	12CA5 epitope of m
13	55	100.0	9	19	AAW76742	Murine signalling
14	55	100.0	9	19	AAW66198	Influenza haemaggl
15	55	100.0	9	19	AAW80131	12CA5 epitope of h
16	55	100.0	9	19	AAW76408	Human PAR-2 12CA5
17	55	100.0	9	19	AAW61471	Influenza haemaggl
18	55	100.0	9	19	AAW50249	Haemagglutinin tag
19	55	100.0	9	19	AAW48632	Chimeric adenovira
20	55	100.0	9	19	AAW60082	Influenza virus ha
21	55	100.0	9	19	AAW60722	Peptide sequence o
22	55	100.0	9	19	AAW59450	Human influenza vi
23	55	100.0	9	19	AAW46333	Binding domain of
24	55	100.0	9	20	AAY31722	Haemagglutinin epi
25	55	100.0	9	20	AAY39621	CTLA-4 VLD CDR loo
26	55	100.0	9	20	AAY26291	FLAG peptide from
27	55	100.0	9	20	AAY06463	Haemagglutinin ant
28	55	100.0	9	20	AAY25085	HA-tag peptide. U
29	55	100.0	9	20	AAW92477	Murine signalling
30	55	100.0	9	20	AAW84206	Peptide comprising
31	55	100.0	9	20	AAW73373	Haemagglutinin epi
32	55	100.0	9	21	AAB29591	HA residues used i
33	55	100.0	9	21	AAB10960	Influenza hemaglut
34	55	100.0	9	21	AAB29434	HA tag, SEQ ID NO:
35	55	100.0	9	21	AAB29033	SNP detection meth
36	55	100.0	9	21	AAB15369	Influenza haemaggl
37	55	100.0	9	21	AAB22878	HA epitope, SEQ ID
38	55	100.0	9	21	AAY96185	Haemagglutinin epi
39	55	100.0	9	21	AAB14332	Influenza virus HA
40	55	100.0	9	21	AAY96884	C-terminal influen
41	55	100.0	9	21	AAB03561	Murine signalling
42	55	100.0	9	21	AAY93541	Amino acid sequenc
43	55	100.0	9	21	AAY87916	Influenza hemagglu
44	55	100.0	9	21	AAY79580	Haemagglutinin epi
45	55	100.0	9	22	AAB61542	12CA5 antibody pep

ALIGNMENTS

RESULT 1
AAR56545
ID AAR56545 standard; Protein; 9 AA.
XX
AC AAR56545;
XX
DT 18-MAR-1995 (first entry)
XX
DE Hal epitope tag.
XX
KW Max-interacting protein; plasmid pJG4-5; fusion protein;
KW cancer; diagnosis; therapy.
XX
OS Homo sapiens.
XX
PN WO9417101-A.
XX
PD 04-AUG-1994.
XX
PF 29-DEC-1993; 93WO-US12643.
XX
PR 29-JAN-1993; 93US-0011398.
XX
PA (GEHO) GEN HOSPITAL CORP.
XX
PI Brent R, Zervos AS;
XX
DR WPI; 1994-264033/32.
XX
PT Max-interacting polypeptide and DNA encoding them - used as
PT anticancer agents and to screen for agents which inhibit cellular
PT proliferation
XX
PS Disclosure; Page 43; 57pp; English.

XX
CC A Max-interacting (Mxi) protein fusion protein with the SV40
CC nuclear localization sequence (AAR56544), the B42 acid blob
CC transcriptional activation domain and the HA1 epitope tag is
CC expressed from pJG4-5 -1. Mxi (and sense/antisense RNA) can be
CC delivered to e.g. bone marrow as therapeutic agents for cancer.
XX
SQ Sequence 9 AA;

Query Match 100.0%; Score 55; DB 15; Length 9;
Best Local Similarity 100.0%; Pred. No. 4.3e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 YPYDVPDYA 9
|||||||||
Db 1 ypydvpdya 9

RESULT 2
AAR78372
ID AAR78372 standard; peptide; 9 AA.
XX
AC AAR78372;
XX
DT 16-APR-1996 (first entry)
XX
DE Viral antigenic epitope of influenza virus haemagglutinin (12CA5).
XX
KW Truncated; tobacco mosaic virus; TMV; movement protein; coat protein;
KW frame shift; mutation; in-frame stop codon; antigen; mammal;
KW viral vaccine; contraceptive; immune response; plant virus; insect.
XX
OS Influenza virus.
XX
PN WO9521248-A1.
XX
PD 10-AUG-1995.
XX
PF 03-FEB-1995; 95WO-US01467.
XX
PR 03-FEB-1994; 94US-0192477.
XX
PA (SCRI) SCRIPPS RES INST.
XX
PI Beachy RN, Fitchen JH;
XX
DR WPI; 1995-283768/37.
XX
PT Overprodn. of heterologous peptide in plants via tobacco mosaic
PT virus infection - in which the coat protein gene is modified by
PT insertion of heterologous sequence, partic. for producing viral
PT antigens for vaccines
XX
PS Claim 10; Page 109; 157pp; English.
XX
CC This sequence represents a viral antigenic epitope of influenza virus
CC haemagglutinin. The DNA sequence encoding this antigen may be included
CC into a modified tobacco mosaic virus (TMV) coat protein gene or a
CC modified movement gene. The modified infectious clone is used to
CC transfect a plant which is then cultured such that the epitope may be
CC recovered from the leaves. The epitope coding sequences are included
CC into the coat protein or movement protein genes such that the translated
CC product contains one or more amino acids immediatly adjacent to the
CC epitope that are readily treated chemically or enzymatically to release
CC the peptide epitope from the precursor protein. Recombinant viruses
CC containing modified movement or coat proteins may be used in the method
CC of the invention for the overproduction of heterologous proteins. The
CC method is esp. used for the production of peptides which cause an
CC antigenic response in mammals , partic. for use in viral vaccines, or
CC for generating a contraceptive immune response. The infectious clones
CC may encode coat proteins for other plant viruses, or proteins conferring
CC protection against insects and diseases.
XX
SQ Sequence 9 AA;

Query Match 100.0%; Score 55; DB 16; Length 9;
Best Local Similarity 100.0%; Pred. No. 4.3e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 YPYDVPDYA 9
|||||||||
Db 1 ypydvpdya 9

RESULT 3
AAW03654
ID AAW03654 standard; Peptide; 9 AA.
XX
AC AAW03654;
XX
DT 22-FEB-1997 (first entry)
XX
DE HA-epitope tag.
XX
KW HA-epitope tag; affinity tail; FADD; Fas receptor; binding;
KW Fas-associating protein with novel death domain; apoptosis;
KW gene therapy; antibody; immunoassay; drug screening; diagnostic;
KW AIDS; antiinflammatory; antitumour; cerebroprotective;
KW neuroprotective.
XX
OS Synthetic.
XX
PN WO9631603-A2.
XX
PD 10-OCT-1996.
XX
PF 28-FEB-1996; 96WO-US02857.
XX
PR 18-MAY-1995; 95US-0443982.
PR 03-APR-1995; 95US-0416379.
XX
PA (UNMI) UNIV MICHIGAN.
XX
PI Dixit VM, O'Rourke K;
XX
DR WPI; 1996-465026/46.
XX
PT FADD protein that binds to cytoplasmic region of Fas receptor - for
PT identifying inhibitors of Fas-associated apoptosis useful for
PT treating e.g. AIDS, leukaemia, stroke, etc
XX
PS Example 1; Page 51; 96pp; English.
XX
CC This HA-epitope tag peptide has been fused to a FADD protein
CC (Fas-associating protein with novel death domain, AAT39397), which
CC binds the cytoplasmic region of a Fas receptor, and modulates
CC apoptosis induced by activation of the receptor by ligand binding.
CC DNA encoding the tag is fused to the FADD gene by PCR, downstream
CC of a cytomegalovirus promoter-enhancer in plasmid pcDNA3HA-FADD.
CC The resulting fusion protein is expressed in Escherichia coli and
CC used in an assay to demonstrate in vivo association of FADD protein
CC and Fas receptor via death domains. FADD DNA may be used in gene
CC therapy, and FADD protein or a corresponding antibody may be used
CC to screen for agents modulating FADD pathway cellular functions
CC and Fas-associated apoptosis, for use in therapy of e.g. AIDS,
CC inflammation, leukaemia, myocardial infarction, degenerative
CC disease, etc.
XX
SQ Sequence 9 AA;

Query Match 100.0%; Score 55; DB 17; Length 9;
Best Local Similarity 100.0%; Pred. No. 4.3e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy      1 YPYDVPDYA 9
          |||||||||
Db      1 ypydvpdya 9


RESULT   4
AAR92029
ID   AAR92029 standard; peptide; 9 AA.
XX
AC   AAR92029;
XX
DT   29-MAY-1996  (first entry)
XX
DE   Hydrophilic peptide for epimorphin modification (1).
XX
KW   Epimorphin; human; mouse; wound; burn; epithelial tissue;
KW   diagnosis; treatment; morphogenetic abnormality; cosmetic;
KW   hair growth stimulator.
XX
OS   Synthetic.
XX
PN   EP698666-A2.
XX
PD   28-FEB-1996.
XX
PF   20-JUN-1995;   95EP-0304270.
XX
PR   31-MAR-1995;   95JP-0099980.
PR   21-JUN-1994;   94JP-0162874.
PR   31-MAR-1995;   95JP-0099979.
XX
PA   (SUME ) SUMITOMO ELECTRIC IND CO.
XX
PI   Hirai Y,  Koshida S,  Oka Y;
XX
DR   WPI; 1996-118213/13.
XX
PT   Novel polypeptide containing an epimorphin functional domain  - has
PT   possible benefits in epithelial tissue treatments, e.g. burns and
PT   for artificial organs
XX
PS   Claim 8; Page 57; 62pp; English.
XX
CC   New polypeptides contain a first portion of 5-99 amino acids joined
CC   to a second portion contg. at least a functional domain of
CC   epimorphin.  The first portion may be selected from the peptides
CC   given in AAR92029 to AAR92036.  The second portion may be full-length
CC   epimorphin (see AAR92037 to AAR92042 for human and mouse epimorphins).
CC   Fragments of epimorphins given in AAT16083 to AAT16090 are used in the
CC   prodn. of modified epimorphins.
CC   The modified epimorphins are useful for the development of diagnosis
CC   and treatment of morphogenetic abnormalities of epithelial tissue or
CC   novel remedies for wounds, eg burns, after surgery and for artificial
CC   organs.  They may also be used as components of cosmetics, hair
CC   growth stimulators, etc.
XX
SQ   Sequence   9 AA;


  Query Match             100.0%;  Score 55;  DB 17;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 4.3e+05;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 YPYDVPDYA 9
          |||||||||
Db      1 ypydvpdya 9


RESULT   5
AAW30296
ID   AAW30296 standard; Peptide; 9 AA.
XX
AC   AAW30296;
XX
DT   14-APR-1998  (first entry)
XX
DE   Epitope for anti-HA tag monoclonal antibody.
XX
KW   Selective target cell activation; G protein-coupled receptor;
KW   RASSL; gene therapy; cell proliferation; haemagglutinin; HA tag;
KW   monoclonal antibody.
XX
OS   Synthetic.
XX
PN   WO9735478-A1.
XX
PD   02-OCT-1997.
XX
PF   25-MAR-1997;   97WO-US05334.
XX
PR   26-MAR-1996;   96US-0622348.
XX
PA   (REGC ) UNIV CALIFORNIA.
XX
PI   Conklin BR;
XX
DR   WPI; 1997-502739/46.
XX
PT   Selective activation of target cell expressing modified G protein
PT   coupled receptor - allows control of cellular proliferation,
PT   especially for amplification of transfected cells in gene therapy
XX
PS   Example 1; Page 57; 117pp; English.
XX
CC   This peptide comprises an epitope for an anti-haemagglutinin (HA)
CC   tag monoclonal antibody.  It can be included in the sequence of a
CC   recombinant G protein-coupled receptor that is activated superiorly
CC   by a synthetic ligand, i.e. a RASSL (see AAW30298 and AAW30299), to
CC   facilitate ease of detection and purification.  RASSLs are used in
CC   a novel method for selectively activating a target cell, providing
CC   direct, selective control of cellular proliferation, especially for
CC   amplification of transfected cells during gene therapy procedures.
XX
SQ   Sequence   9 AA;


  Query Match             100.0%;  Score 55;  DB 18;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 4.3e+05;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 YPYDVPDYA 9
          |||||||||
Db      1 ypydvpdya 9


RESULT   6
AAW27094
ID   AAW27094 standard; peptide; 9 AA.
XX
AC   AAW27094;
XX
DT   19-NOV-1997  (first entry)
XX
DE   Transforming growth factor-beta activated kinase epitope HA-TAK1.
XX
KW   TGF-beta; signal transmission; TGF-beta activated kinase;
KW   MAPK kinase activator; AMK-1; bone morphogenetic protein; BMP;
KW   protein kinase; epitope.
XX
OS   Synthetic.
XX
PN   JP09163990-A.
XX
PD   24-JUN-1997.
XX
```

```
PF   27-SEP-1996;   96JP-0256747.
XX
PR   24-JUL-1996;   96US-0685625.
PR   29-SEP-1995;   95JP-0253549.
XX
PA   (CHUS ) CHUGAI PHARM CO LTD.
PA   (UENO/) UENO N.
XX
DR   WPI; 1997-380171/35.
XX
PT   DNA encoding transforming growth factor-beta-activated kinase, TAK-1
PT   - useful for studying the TGF-beta signal transmission system
XX
PS   Example 4; Page 8; 20pp; Japanese.
XX
CC   The present sequence represents an epitope, HA-TAK1, of transforming
CC   growth factor-beta (TGF-beta) activated kinase, TAK-1. The DNA
CC   encoding TAK-1 is used to produce the TAK-1 protein which is involved
CC   in the TGF-beta family signal transmission system. TAK-1, also known
CC   as activator of MAPK Kinase (AMK-1), is an enzyme which is activated
CC   by TGF-beta and bone morphogenetic protein (BMP) and activates MAPK
CC   kinase by phosphorylation.
XX
SQ   Sequence   9 AA;


  Query Match             100.0%;  Score 55;  DB 18;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 4.3e+05;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 YPYDVPDYA 9
           |||||||||
Db       1 ypydvpdya 9


RESULT   7
AAW00970
ID   AAW00970 standard; Peptide; 9 AA.
XX
AC   AAW00970;
XX
DT   12-NOV-1997  (first entry)
XX
DE   Haemagglutinin epitope.
XX
KW   DNA binding protein; RNA binding protein; amphipathic peptide;
KW   acidic extension peptide; gene control; gene regulation;
KW   transcription; dominant negative protein; cancer; drug therapy;
KW   drug design; EBP; leucine zipper; haemagglutinin.
XX
OS   Synthetic.
XX
PN   WO9705249-A2.
XX
PD   13-FEB-1997.
XX
PF   31-JUL-1996;   96WO-US12590.
XX
PR   31-JUL-1996;   96US-0001654.
PR   31-JUL-1995;   95US-0001654.
PR   29-MAY-1996;   96US-0018496.
XX
PA   (KRYL/) KRYLOV D.
PA   (USSH ) US DEPT HEALTH & HUMAN SERVICES.
PA   (VINS/) VINSON C R.
XX
PI   Krylov D,  Vinson CR;
XX
DR   WPI; 1997-145687/13.
XX
PT   New nucleic acid binding proteins - having an acidic amino acid
PT   sequence extension at the amino-terminus, to increase ability to
PT   regulate gene transcription, useful e.g. in cancer therapeutics
XX
PS   Claim 16; Page 58; 144pp; English.
XX
CC   This peptide sequence comprises a haemagglutinin epitope that can
CC   be attached to the N-terminus of a nucleic acid binding protein
CC   (NABP) such as a leucine zipper bZIP protein or bHLH protein.
CC   Claimed NABPs having an appended acidic extension peptide (see
CC   AAW00958-65) can regulate the function of a target nucleic acid or
CC   gene to which they are bound, and act as potent dominant-negative
CC   regulators of gene transcription, cell growth and cell proliferation.
CC   They can be used in cancer therapeutics, to treat diseases caused by
CC   eukaryotic microorganisms or by viruses, and as tools for drug
CC   development, rational drug design, and drug and gene therapies.
XX
SQ   Sequence   9 AA;


  Query Match             100.0%;  Score 55;  DB 18;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 4.3e+05;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 YPYDVPDYA 9
           |||||||||
Db       1 ypydvpdya 9


RESULT   8
AAW23660
ID   AAW23660 standard; Peptide; 9 AA.
XX
AC   AAW23660;
XX
DT   10-OCT-1997  (first entry)
XX
DE   pJG4-5 library plasmid HA1 epitope tag.
XX
KW   Human; cell differentiation; survival; carcinoma; psoriasis;
KW   ubiquitination; p53; tumour suppressor; homeostasis; papilloma virus;
KW   epithelial cell; acne; icthyosis; aphthous ulcer; hair growth;
KW   antibody; cell proliferation.
XX
OS   Synthetic.
XX
PN   WO9640767-A2.
XX
PD   19-DEC-1996.
XX
PF   06-JUN-1996;   96WO-US09040.
XX
PR   07-JUN-1995;   95US-0484878.
XX
PA   (MITO-) MITOTIX INC.
XX
PI   Beer-Romero PL,  Draetta G,  Rolfe M;
XX
DR   WPI; 1997-087053/08.
XX
PT   E6AP-binding proteins and related nucleic acid - useful for
PT   modulating cell differentiation, survival etc., partic. for
PT   treatment and diagnosis of carcinoma, psoriasis, etc
XX
PS   Disclosure; Page 73; 83pp; English.
XX
CC   The present sequence represents the HA1 epitope tag of the pJG4-5
CC   library plasmid used during the isolation of the E6AP-binding proteins.
CC   E6AP mediates ubiquitination and so the inactivation of e.g. p53 tumour
CC   suppressor. The E6AP-binding protein is likely to be involved in normal
CC   cell homeostasis and in the pathogenesis of proliferative and
CC   differentiation disorders, e.g. regulation of gene expression or the
CC   cell cycle, modification of cell surface receptors, biogenesis of
CC   ribosomes and DNA repair. The protein, which can optionally be
CC   generated in vivo by gene therapy, may useful in treatment and
CC   prevention of papilloma virus infected transformed cells and carcinoma,
```

CC and may also be used to regulate epithelial cell processes more
CC generally, e.g. in cases of psoriasis, acne, icthyosis, aphthous
CC ulcers. It may also be used to control wound healing, and inhibit growth
CC of hair. It can also be used to generate antibodies which are used in
CC immunoassays to determine the protein levels. The encoding nucleic acid
CC can be used to prepare recombinant proteins and oligonucleotides useful
CC as probes and primers for diagnostic detection of (mutant) mRNA for the
CC protein in (transformed) cells, also for antisense therapy. It can also
CC be used for the detection of mutations in E6AP-binding protein-encoding
CC genes, mis-expression of these genes or mis-incorporation of the protein
CC in a transcription regulatory complex containing E6AP can be used to
CC assess risk of disorders characterised by cell proliferation.
XX
SQ Sequence 9 AA;

Query Match 100.0%; Score 55; DB 18; Length 9;
Best Local Similarity 100.0%; Pred. No. 4.3e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 YPYDVPDYA 9
|||||||||
Db 1 ypydvpdya 9

RESULT 9
AAW08162
ID AAW08162 standard; Peptide; 9 AA.
XX
AC AAW08162;
XX
DT 26-AUG-1997 (first entry)
XX
DE Purification tag of a TGF-beta fusion protein.
XX
KW Transforming growth factor-beta fusion protein; wound healing;
KW artificial skin; surgery recovery time.
XX
OS Synthetic.
XX
PN WO9639430-A1.
XX
PD 12-DEC-1996.
XX
PF 05-JUN-1996; 96WO-US08973.
XX
PR 06-JUN-1995; 95US-0470837.
XX
PA (CHEU/) CHEUNG D T.
PA (HALL/) HALL F L.
PA (NIMN/) NIMNI M E.
PA (TUAN/) TUAN T.
PA (WULL/) WU L.
XX
PI Cheung DT, Hall FL, Nimni ME, Tuan T, Wu L;
XX
DR WPI; 1997-043065/04.
DR N-PSDB; AAT42766.
XX
PT Prepn. of transforming growth factor-beta fusion protein - useful to
PT reduce surgery recovery time and to prepare artificial skin
XX
PS Claim 5; Page 44; 59pp; English.
XX
CC A novel transforming growth factor-beta (TGF-beta) fusion protein
CC comprises a purification tag and a TGF active fragment. The present
CC sequence represents a specifically claimed purification tag.
CC Additionally, the fusion protein may comprise proteinase-sensitive
CC linker sites and binding domain so the protein sequence may contain
CC some or all of the following elements: purification tag:proteinase
CC site:ECM binding site:proteinase site:TGF-beta. TGF-beta promotes
CC wound healing, and the fusion protein can be used to reduce surgery
CC recovery time and in the preparation of artificial skin. The inclusion
CC of a purification tag facilitates purification of the fusion protein.
CC The proteinase site is included to permit cleavage and release of the
CC purification tag after purification if desired. The extracellular
CC matrix binding site facilitates delivery of the fusion protein to the
CC desired site of action. Delivery of the TGF-beta to the site to be
CC treated reduces the amount of TGF-beta required to be administered to
CC be effective and reduces the concentration of circulating TGF-beta
CC which may result in undesirable effects.
XX
SQ Sequence 9 AA;

Query Match 100.0%; Score 55; DB 18; Length 9;
Best Local Similarity 100.0%; Pred. No. 4.3e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 YPYDVPDYA 9
|||||||||
Db 1 ypydvpdya 9

RESULT 10
AAW12385
ID AAW12385 standard; Peptide; 9 AA.
XX
AC AAW12385;
XX
DT 18-JUN-1997 (first entry)
XX
DE Influenza virus haemagglutinin epitope tag.
XX
KW Plasmid pPhOx.TM; eukaryotic expression vector; transfected cell;
KW single chain antibody; sFv; molecular hook; haemagglutinin;
KW epitope tag.
XX
OS Influenza virus.
XX
PN WO9708186-A1.
XX
PD 06-MAR-1997.
XX
PF 23-AUG-1996; 96WO-US15819.
XX
PR 24-AUG-1995; 95US-0518835.
XX
PA (INVI-) INVITROGEN CORP.
XX
PI Baytan A, Bernhard A, Chang M, Chesnut RD, Hoeffler JP;
XX
DR WPI; 1997-179169/16.
XX
PT Eukaryotic expression vector - for identification and separation of
PT transfected cells from total cell population
XX
PS Example 1; Page 27; 83pp; English.
XX
CC A DNA sequence encoding an influenza virus haemagglutinin (HA)
CC epitope tag (AAW12385) and the murine Ig kappa-chain V-J2-C region
CC signal peptide (AAW12384) was subcloned from a single chain antibody
CC (sFv) contg. vector (pCR3.2) as an EcoRV-SalI fragment and ligated
CC with EcoRV/SalI-digested pCR3.1.1. The resulting vector was used
CC in the construction of eukaryotic expression vector pPhOx.TM (see
CC also AAT63235). Expression of the HA epitope tag allowed detection
CC of the expressed sFv by monoclonal antibody 12CA5. Vector pPhOx.TM
CC is used in novel methods for the identification and isolation of
CC transfected eukaryotic cells.
XX
SQ Sequence 9 AA;

Query Match 100.0%; Score 55; DB 18; Length 9;
Best Local Similarity 100.0%; Pred. No. 4.3e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy      1 YPYDVPDYA 9
          |||||||||
Db      1 ypydvpdya 9
```

RESULT 11
AAW01661
ID AAW01661 standard; Peptide; 9 AA.
XX
AC AAW01661;
XX
DT 25-MAR-1997 (first entry)
XX
DE 10-mer peptide tag derived from Influenza haemagglutinin.
XX
KW monoclonal; recombinant; antibody; bispecific; antigenic site;
KW Fab fragment; complementary interactive domain; leucine zipper;
KW c-fos; c-jun; treatment; cancer; viral infection; primer; PCR;
KW polymerase chain reaction; amplification; peptide tag.
XX
OS Synthetic.
XX
PN US5582996-A.
XX
PD 10-DEC-1996.
XX
PF 04-DEC-1990; 90US-0622983.
XX
PR 31-JUL-1992; 92US-0917074.
PR 04-DEC-1990; 90US-0622983.
PR 03-DEC-1991; 91WO-US09019.
PR 27-MAY-1994; 94US-0250656.
XX
PA (WIST-) WISTAR INST ANATOMY & BIOLOGY.
XX
PI Curtis PJ;
XX
DR WPI; 1997-042302/04.
XX
PT Bi:specific antibodies comprising two Fab fragments linked through
PT leucine zipper - useful to detect selected antigen or to treat
PT cancer and viral infections
XX
PS Example 1; Column 10; 10pp; English.
XX
CC Recombinant antibodies capable of binding to two different antigenic
CC sites, contain Fab fragments from the same or, pref., different
CC antibodies, which are brought into association by complementary
CC interactive domains, esp. leucine zipper domains, which have been
CC inserted into a region of the antibody heavy chain constant region.
CC Recombinant proteins c-fos LZ/Fab OKT2 and c-jun LZ/Fab OKT3 showed
CC little or no binding to normal lymphocytes when analysed individually
CC but bound very clearly to the cells when mixed to form a homodimeric
CC antibody. The labelled antibodies are useful for detecting a selected
CC antigen (bound by the first or second Fab) in a biological sample or
CC treatment of cancer and viral infections. AAW01661 is a 10-mer peptide
CC tag derived from Influenza haemagglutinin. DNA encoding the 10-mer was
CC cloned into lambda Hc2 to provide a peptide tag which will be present
CC on the heavy chain cloned into pHc2.
XX
SQ Sequence 9 AA;

```
  Query Match             100.0%;  Score 55;  DB 18;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 4.3e+05;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 YPYDVPDYA 9
          |||||||||
Db      1 ypydvpdya 9
```

RESULT 12
AAW81556
ID AAW81556 standard; peptide; 9 AA.
XX
AC AAW81556;
XX
DT 04-FEB-1999 (first entry)
XX
DE 12CA5 epitope of mouse LIFR alpha-chain.
XX
KW LIFR; leukaemia inhibitory factor; LIF; interspecies; chimeric;
KW haemopoietin domain; immunoglobulin; cytokine; interleukin;
KW colony-stimulating factor; mouse; epitope.
XX
OS Mus sp.
XX
PN WO9848011-A1.
XX
PD 29-OCT-1998.
XX
PF 21-APR-1998; 98WO-AU00282.
XX
PR 21-APR-1997; 97AU-0006328.
XX
PA (HALL-) HALL INST MEDICAL RES WALTER & ELIZA.
XX
PI Layton MJ, Metcalf D, Nicola NA, Owczarek CM, Zhang Y;
XX
DR WPI; 1998-609986/51.
XX
PT New polypeptide containing haemopoietin domain and
PT immunoglobulin-like domain - also related nucleic acid, binds
PT cytokines, specifically leukaemia inhibitory factor, and is used to
PT treat inflammation
XX
PS Example 1; Page 24; 58pp; English.
XX
CC This represents an in-frame 12CA5 peptide epitope of the alpha chain of
CC mouse leukaemia inhibitory factor (LIF) receptor (LIFR). The invention
CC relates to interspecies LIFR alpha-chain chimeras. The chimeric
CC polypeptide comprises (i) a haemopoietin domain, linked, bound or
CC associated with (ii) an immunoglobulin (Ig)-like domain, such that the
CC chimera can bind cytokines. The chimeric polypeptide specifically binds
CC LIF, so is used to treat conditions that require antagonism of LIF,
CC specifically inflammation, but more generally may be designed to bind
CC other cytokines such as interleukins and colony-stimulating factors.
XX
SQ Sequence 9 AA;

```
  Query Match             100.0%;  Score 55;  DB 19;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 4.3e+05;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 YPYDVPDYA 9
          |||||||||
Db      1 ypydvpdya 9
```

RESULT 13
AAW76742
ID AAW76742 standard; Protein; 9 AA.
XX
AC AAW76742;
XX
DT 15-JAN-1999 (first entry)
XX
DE Murine signalling chimeric protein peptide #7.
XX
KW Induction; intracellular aggregation; cytoplasmic tail; zeta chain;
KW T cell receptor/CD3 complex; signal; transcription; reporter gene;
KW homodimerization; Fas receptor; cell-specific apoptosis; Gal14; VP16;
KW heterodimerization; primer.

XX
OS Mus sp.
XX
PN US5830462-A.
XX
PD 03-NOV-1998.
XX
PF 07-JUN-1995; 95US-0478386.
XX
PR 14-FEB-1995; 95US-0388653.
PR 12-FEB-1993; 93US-0017931.
PR 16-JUL-1993; 93US-0092977.
PR 07-JAN-1994; 94US-0179748.
PR 11-FEB-1994; 94US-0196043.
PR 18-AUG-1994; 94US-0292597.
PR 07-JUN-1995; 95US-0478386.
XX
PA (HARD) HARVARD COLLEGE.
PA (STRD) UNIV LELAND S STANFORD.
XX
PI Belshaw P, Crabtree GR, Schreiber SL, Spencer DM;
PI Wandless TJ;
XX
DR WPI; 1998-609192/51.
DR N-PSDB; AAV62953.
XX
PT Inducing a biological process in a cell using ligand-induced gene
PT expression for e.g. cell death - using DNA construct encoding
PT ligand-binding domain and dimerisable protein domain
XX
PS Disclosure; Column 71-72; 95pp; English.
XX
CC AAW76736-W76764 are peptides used to design primers used in a novel
CC method for inducing a biological process in a cell. The method can be
CC used to induce the intracellular aggregation of the cytoplasmic tail of
CC the zeta chain of the T cell receptor/CD3 complex, leading to signalling
CC and transcription of a reporter gene. The method also has applications
CC for the homodimerization of the cytoplasmic tail of the Fas receptor,
CC leading to cell-specific apoptosis and for the heterodimerization of Gal4
CC and VP16, leading to direct transcription of a reporter gene.
XX
SQ Sequence 9 AA;

Query Match 100.0%; Score 55; DB 19; Length 9;
Best Local Similarity 100.0%; Pred. No. 4.3e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 YPYDVPDYA 9
|||||||||
Db 1 ypydvpdya 9

RESULT 14
AAW66198
ID AAW66198 standard; Peptide; 9 AA.
XX
AC AAW66198;
XX
DT 21-DEC-1998 (first entry)
XX
DE Influenza haemagglutinin epitope.
XX
KW Intein; recA; DNA repair enzyme; vaccine; vector; antigen;
KW open reading frame; haemagglutinin.
XX
OS Influenza virus.
XX
PN WO9840394-A1.
XX
PD 17-SEP-1998.
XX
PF 10-MAR-1998; 98WO-US04805.
XX
PR 13-MAR-1997; 97US-0816721.
XX
PA (YESH) UNIV YESHIVA EINSTEIN COLLEGE.
XX
PI Daugelat S, Jacobs WR;
XX
DR WPI; 1998-520797/44.
XX
PT New vector constructs for selection of open reading frames -
PT comprise origin of replication and DNA sequences encoding selectable
PT marker, intein and regulatory elements for expression
XX
PS Disclosure; Fig 7; 83pp; English.
XX
CC This peptide comprises the influenza virus haemagglutinin epitope.
CC A synthetic oligonucleotide (see AAV07684) encoding the epitope has
CC been inserted into plasmid pYUB763 (see AAV07706), which contains DNA
CC encoding Mycobacterium tuberculosis recA intein (see also AAV07705)
CC inserted into an aph selectable marker gene and includes an
CC internal BglII site. E. coli transformants showed a high splicing
CC frequency. The invention provides novel vector constructs that
CC include a DNA sequence encoding an intein inserted into a
CC selectable marker gene such as aph. DNA of interest can be cloned
CC into a unique restriction site of the intein, and the vector used
CC as a vaccine or transformed into a vaccine vector. The vectors may
CC also be used in methods of selecting translated open reading frames
CC or genes, leading to the identification of protective antigens of
CC pathogenic organisms.
XX
SQ Sequence 9 AA;

Query Match 100.0%; Score 55; DB 19; Length 9;
Best Local Similarity 100.0%; Pred. No. 4.3e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 YPYDVPDYA 9
|||||||||
Db 1 ypydvpdya 9

RESULT 15
AAW80131
ID AAW80131 standard; peptide; 9 AA.
XX
AC AAW80131;
XX
DT 15-DEC-1998 (first entry)
XX
DE 12CA5 epitope of haemagglutinin.
XX
KW Haemagglutinin 12CA5 epitope; polyclonal rabbit antibody HA.11;
KW human beta1-adrenergic receptor; beta1-AR; transgene;
KW heart tissue-specific promoter; transgenic animal model;
KW heart muscle disease; heart failure.
XX
OS Synthetic.
XX
PN WO9844092-A1.
XX
PD 08-OCT-1998.
XX
PF 02-APR-1998; 98WO-US06791.
XX
PR 03-APR-1997; 97US-0041966.
XX
PA (UYTE-) UNIV TECHNOLOGY CORP.
XX
PI Bristow MR, Port JD;
XX
DR WPI; 1998-557104/47.
XX

PT Transgenic mice as models for heart disease - having incorporated in
PT their genome a heart tissue-specific promoter operatively linked to
PT DNA coding for a beta-1-adrenergic receptor
XX
PS Example 1; Page 15; 40pp; English.
XX
CC The present sequence represents the 12CA5 epitope of haemagglutinin. It
CC is recognised by the polyclonal rabbit antibody HA.11. The sequence was
CC produced by annealing 2 primers encoding the present sequence. These
CC were used to create a epitope tag at the end of human beta1-adrenergic
CC receptor (beta1-AR) cDNA, which is used to create the transgene of the
CC invention. The specification describes a transgenic animal, especially
CC a mouse, which has incorporated into its genome a transgene comprising
CC a heart tissue-specific promoter operatively linked to a coding
CC sequence comprising beta1-1AR, the transgene being expressed in at least
CC the myocardium of the heart of the transgenic animal. The transgenic
CC animal can be used as a model for heart muscle disease and heart failure
CC in a mammal. The transgene can also be used for treating heart failure.
XX
SQ Sequence 9 AA;

Query Match 100.0%; Score 55; DB 19; Length 9;
Best Local Similarity 100.0%; Pred. No. 4.3e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 YPYDVPDYA 9
|||||||||
Db 1 ypydvpdya 9

Search completed: April 11, 2002, 10:15:43
Job time: 93 sec

GenCore version 4.5

OM protein - protein search, using sw model

Run on: April 11, 2002, 10:14:10 ; Search time 24.73 Seconds
(without alignments)
8.190 Million cell updates/sec

Title: US-09-284-787-1
Perfect score: 55
Sequence: 1 YPYDVPDYA 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 212252 seqs, 22503292 residues

Total number of hits satisfying chosen parameters: 212252

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	55	100.0	9	1	US-08-011-398B-6	Sequence 6, Appli
2	55	100.0	9	1	US-08-090-148-2	Sequence 2, Appli
3	55	100.0	9	1	US-08-370-225-6	Sequence 6, Appli
4	55	100.0	9	1	US-08-167-982-1	Sequence 1, Appli
5	55	100.0	9	1	US-08-634-060-31	Sequence 31, Appl
6	55	100.0	9	1	US-08-464-051-6	Sequence 6, Appli
7	55	100.0	9	1	US-08-461-859-6	Sequence 6, Appli
8	55	100.0	9	1	US-08-470-837-28	Sequence 28, Appl
9	55	100.0	9	2	US-08-478-386A-19	Sequence 19, Appl
10	55	100.0	9	2	US-08-292-597-19	Sequence 19, Appl
11	55	100.0	9	2	US-08-462-498-6	Sequence 6, Appli
12	55	100.0	9	2	US-08-922-267A-60	Sequence 60, Appl
13	55	100.0	9	2	US-08-388-653-19	Sequence 19, Appl
14	55	100.0	9	2	US-08-473-985-19	Sequence 19, Appl
15	55	100.0	9	2	US-08-690-011A-17	Sequence 17, Appl
16	55	100.0	9	2	US-08-690-011A-49	Sequence 49, Appl
17	55	100.0	9	2	US-08-685-625A-4	Sequence 4, Appli
18	55	100.0	9	2	US-08-687-559-14	Sequence 14, Appl
19	55	100.0	9	2	US-08-700-846-11	Sequence 11, Appl
20	55	100.0	9	2	US-08-287-537B-3	Sequence 3, Appli
21	55	100.0	9	2	US-08-483-898-19	Sequence 19, Appl
22	55	100.0	9	3	US-09-087-716-19	Sequence 19, Appl
23	55	100.0	9	3	US-08-853-733B-28	Sequence 28, Appl
24	55	100.0	9	3	US-08-554-385-6	Sequence 6, Appli
25	55	100.0	9	3	US-08-518-835-6	Sequence 6, Appli
26	55	100.0	9	3	US-09-157-753-19	Sequence 19, Appl
27	55	100.0	9	3	US-09-157-230-19	Sequence 19, Appl
28	55	100.0	9	3	US-09-087-811-19	Sequence 19, Appl
29	55	100.0	9	3	US-08-968-747-12	Sequence 12, Appl
30	55	100.0	9	3	US-09-156-855-19	Sequence 19, Appl
31	55	100.0	9	3	US-09-030-613-36	Sequence 36, Appl
32	55	100.0	9	3	US-08-493-071-21	Sequence 21, Appl
33	55	100.0	9	4	US-09-158-010-19	Sequence 19, Appl
34	55	100.0	9	4	US-08-360-694-3	Sequence 3, Appli
35	55	100.0	9	4	US-09-087-647-19	Sequence 19, Appl
36	55	100.0	9	4	US-09-284-033-17	Sequence 17, Appl
37	55	100.0	9	4	US-09-053-293-1	Sequence 1, Appli
38	55	100.0	9	4	US-09-208-966-17	Sequence 17, Appl
39	55	100.0	9	4	US-09-100-802-17	Sequence 17, Appl
40	55	100.0	9	4	US-09-140-084-1	Sequence 1, Appli
41	55	100.0	9	4	US-09-382-950-8	Sequence 8, Appli
42	55	100.0	9	4	US-09-451-905-36	Sequence 36, Appl
43	55	100.0	9	4	US-09-382-736B-9	Sequence 9, Appli
44	55	100.0	9	5	PCT-US93-10069-6	Sequence 6, Appli
45	55	100.0	9	5	PCT-US94-10936-3	Sequence 3, Appli

ALIGNMENTS

```
RESULT   1
US-08-011-398B-6
; Sequence 6, Application US/08011398B
; Patent No. 5512473
;  GENERAL INFORMATION:
;    APPLICANT:  Roger Brent
;    APPLICANT:  Antonis S. Zervos
;    TITLE OF INVENTION:  MAX-INTERACTING PROTEINS AND RELATED
;    TITLE OF INVENTION:  MOLECULES AND METHODS
;    NUMBER OF SEQUENCES:  20
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Fish & Richardson
;      STREET:  225 Franklin Street
;      CITY:  Boston
;      STATE:  Massachusetts
;      COUNTRY:  U.S.A.
;      ZIP:  02110-2804
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5" Diskette, 1.44  Mb
;      COMPUTER:  IBM PS/2 Model 50Z or 55SX
;      OPERATING SYSTEM:  MS-DOS (Version 5.0)
;      SOFTWARE:  WordPerfect (Version 5.1)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/011,398B
;      FILING DATE:  29 JAN 1993
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Paul T. Clark
;      REGISTRATION NUMBER:  30,162
;      REFERENCE/DOCKET NUMBER:  00786/160001
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617) 542-5070
;      TELEFAX:  (617) 542-8906
;      TELEX:  200154
;  INFORMATION FOR SEQ ID NO:  6:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  9
;      TYPE:  amino acid
;      STRANDEDNESS:  N/A
;      TOPOLOGY:  linear
US-08-011-398B-6


  Query Match           100.0%;  Score 55;  DB 1;  Length 9;
  Best Local Similarity  100.0%;  Pred. No. 1.6e+05;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;
```

```
Qy      1 YPYDVPDYA 9
          |||||||||
Db      1 YPYDVPDYA 9


RESULT   2
US-08-090-148-2
; Sequence 2, Application US/08090148
; Patent No. 5534257
;  GENERAL INFORMATION:
;    APPLICANT:  Mastico, Robert Allan
;    APPLICANT:  Stockley, Peter George
;    APPLICANT:  Talbot, Simon John
;    TITLE OF INVENTION:  Antigen-Presenting Capsid with
;    TITLE OF INVENTION:  Fusion MS2-Coat Protein
;    NUMBER OF SEQUENCES:  6
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Rosenman & Colin
;      STREET:  575 Madison Avenue
;      CITY:  New York
;      STATE:  NY
;      COUNTRY:  U.S.A.
;      ZIP:  10022-2585
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 3.5", 1.44Mb
;      COMPUTER:  IBM PS2-486
;      OPERATING SYSTEM:  MS-DOS
;      SOFTWARE:  WordPerfect 5.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/090,148
;      FILING DATE:  08/11/93
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  GB  9101550.3
;      FILING DATE:  01/24/91
;      APPLICATION NUMBER:  PCT/GB92/00124
;      FILING DATE:  01/22/92
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Nissenbaum, Israel
;      REGISTRATION NUMBER:  27,582
;      REFERENCE/DOCKET NUMBER:
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212) 940-8636
;      TELEFAX:  (212) 940-6404
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  9 AMINO ACIDS
;      TYPE:  AMINO ACID
;      TOPOLOGY:  NOT RELEVANT
;    MOLECULE TYPE:  PEPTIDE
US-08-090-148-2


  Query Match           100.0%;  Score 55;  DB 1;  Length 9;
  Best Local Similarity  100.0%;  Pred. No. 1.6e+05;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 YPYDVPDYA 9
          |||||||||
Db      1 YPYDVPDYA 9


RESULT   3
US-08-370-225-6
; Sequence 6, Application US/08370225
; Patent No. 5580736
;  GENERAL INFORMATION:
;    APPLICANT:  Brent, Roger
;    APPLICANT:  Gyuris, Jeno
;    APPLICANT:  Golemis, Erica
;    TITLE OF INVENTION:  Interaction Trap System for Isolating
;    TITLE OF INVENTION:  No. 5580736el Proteins
;    NUMBER OF SEQUENCES:  33
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Fish & Richardson
;      STREET:  225 Franklin Street
;      CITY:  Boston
;      STATE:  Massachusetts
;      COUNTRY:  U.S.A.
;      ZIP:  02110-2804
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5" Diskette, 1.44 Mb
;      COMPUTER:  IBM PS/2 Model 50Z or 55SX
;      OPERATING SYSTEM:  MS-DOS (Version 5.0)
;      SOFTWARE:  WordPerfect (Version 5.1)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/370,225
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/969,038
;      FILING DATE:  10/30/92
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Clark, Paul T.
;      REGISTRATION NUMBER:  30,162
;      REFERENCE/DOCKET NUMBER:  00786/143001
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617) 542-5070
;      TELEFAX:  (617) 542-8906
;      TELEX:  200154
;  INFORMATION FOR SEQ ID NO:  6:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  9
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  linear
US-08-370-225-6


  Query Match           100.0%;  Score 55;  DB 1;  Length 9;
  Best Local Similarity  100.0%;  Pred. No. 1.6e+05;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 YPYDVPDYA 9
          |||||||||
Db      1 YPYDVPDYA 9


RESULT   4
US-08-167-982-1
; Sequence 1, Application US/08167982
; Patent No. 5698424
;  GENERAL INFORMATION:
;    APPLICANT:  Mastico, Robert Allan
;    APPLICANT:  Stockley, Peter George
;    TITLE OF INVENTION:  CAPSID FORMING AND CYSTEINE
;    TITLE OF INVENTION:  MODIFIED CHIMAERIC MS-2-COAT
;    TITLE OF INVENTION:  PROTEIN
;    NUMBER OF SEQUENCES:  1
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Graham & James LLP
;      STREET:  885 Third Avenue
;      CITY:  New York
;      STATE:  NY
;      COUNTRY:  U.S.A.
;      ZIP:  10022-4834
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 3.5", 1.44Mb
;      COMPUTER:  IBM Pentium
;      OPERATING SYSTEM:  MS-DOS
;      SOFTWARE:  WordPerfect 5.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/167,982
;      FILING DATE:  01/03/95
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
```

```
;      APPLICATION NUMBER:  UK  9114003.8
;      FILING DATE:  06/28/91
;      APPLICATION NUMBER:  PCT/GB92/01159
;      FILING DATE:  06/26/92
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Reingold, Jesse
;      REGISTRATION NUMBER:  20,461
;      REFERENCE/DOCKET NUMBER:  Mastico
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212) 848-1032
;      TELEFAX:  (212) 688-2449
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  9 AMINO ACIDS
;      TYPE:  AMINO ACID
;      TOPOLOGY:  NOT RELEVANT
;    MOLECULE TYPE:  PEPTIDE
US-08-167-982-1


  Query Match           100.0%;  Score 55;  DB 1;  Length 9;
  Best Local Similarity  100.0%;  Pred. No. 1.6e+05;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 YPYDVPDYA 9
          |||||||||
Db      1 YPYDVPDYA 9


RESULT   5
US-08-634-060-31
; Sequence 31, Application US/08634060
; Patent No. 5712136
;  GENERAL INFORMATION:
;    APPLICANT:  Wickham, Thomas J.
;    APPLICANT:  Kovesdi, Imre
;    APPLICANT:  Roelvink, Petrus W.
;    TITLE OF INVENTION:  ADENOVIRAL-MEDIATED CELL TARGETING COMMANDED BY
;    TITLE OF INVENTION:  THE ADENOVIRUS PENTON BASE PROTEIN
;    NUMBER OF SEQUENCES:  60
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Leydig, Voit & Mayer, Ltd.
;      STREET:  Two Prudential Plaza, Suite 4900
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  USA
;      ZIP:  60601
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk.
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/634,060
;      FILING DATE:
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/303,162
;      FILING DATE:  08-SEP-1994
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Kilyk, John Jr.
;      REGISTRATION NUMBER:  30763
;      REFERENCE/DOCKET NUMBER:  71602
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (312) 616-5600
;      TELEFAX:  (312) 616-5700
;  INFORMATION FOR SEQ ID NO:  31:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  9 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  unknown
;    MOLECULE TYPE:  peptide
US-08-634-060-31


  Query Match           100.0%;  Score 55;  DB 1;  Length 9;
  Best Local Similarity  100.0%;  Pred. No. 1.6e+05;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 YPYDVPDYA 9
          |||||||||
Db      1 YPYDVPDYA 9


RESULT   6
US-08-464-051-6
; Sequence 6, Application US/08464051
; Patent No. 5780262
;  GENERAL INFORMATION:
;    APPLICANT:  Roger Brent
;    APPLICANT:  Antonis S. Zervos
;    TITLE OF INVENTION:  MAX-INTERACTING PROTEINS AND RELATED
;    TITLE OF INVENTION:  MOLECULES AND METHODS
;    NUMBER OF SEQUENCES:  20
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Fish & Richardson
;      STREET:  225 Franklin Street
;      CITY:  Boston
;      STATE:  Massachusetts
;      COUNTRY:  U.S.A.
;      ZIP:  02110-2804
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5" Diskette, 1.44  Mb
;      COMPUTER:  IBM PS/2 Model 50Z or 55SX
;      OPERATING SYSTEM:  MS-DOS (Version 5.0)
;      SOFTWARE:  WordPerfect (Version 5.1)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/464,051
;      FILING DATE:  05 JUN 1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/011,398
;      FILING DATE:  29 JAN 1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Paul T. Clark
;      REGISTRATION NUMBER:  30,162
;      REFERENCE/DOCKET NUMBER:  00786/160002
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617) 542-5070
;      TELEFAX:  (617) 542-8906
;      TELEX:  200154
;  INFORMATION FOR SEQ ID NO:  6:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  9
;      TYPE:  amino acid
;      STRANDEDNESS:  N/A
;      TOPOLOGY:  linear
US-08-464-051-6


  Query Match           100.0%;  Score 55;  DB 1;  Length 9;
  Best Local Similarity  100.0%;  Pred. No. 1.6e+05;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 YPYDVPDYA 9
          |||||||||
Db      1 YPYDVPDYA 9


RESULT   7
US-08-461-859-6
; Sequence 6, Application US/08461859
; Patent No. 5786169
;  GENERAL INFORMATION:
```

```
;    APPLICANT:  Brent, Roger
;    APPLICANT:  Gyuris, Jeno
;    APPLICANT:  Golemis, Erica
;    TITLE OF INVENTION:  Interaction Trap System for Isolating
;    TITLE OF INVENTION:  No. 5786169el Proteins
;    NUMBER OF SEQUENCES:  35
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Fish & Richardson P.C.
;      STREET:  225 Franklin Street
;      CITY:  Boston
;      STATE:  Massachusetts
;      COUNTRY:  U.S.A.
;      ZIP:  02110-2804
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5" Diskette, 1.44 Mb
;      COMPUTER:  IBM PS/2 Model 50Z or 55SX
;      OPERATING SYSTEM:  MS-DOS (Version 5.0)
;      SOFTWARE:  WordPerfect (Version 5.1)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/461,859
;      FILING DATE:  June 5, 1995
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/370,225
;      FILING DATE:  January 9, 1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/969,038
;      FILING DATE:  October 30, 1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Lech, Karen F.
;      REGISTRATION NUMBER:  35,238
;      REFERENCE/DOCKET NUMBER:  00786/143002
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617) 542-5070
;      TELEFAX:  (617) 542-8906
;      TELEX:  200154
;  INFORMATION FOR SEQ ID NO:  6:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  9
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  linear
US-08-461-859-6


  Query Match           100.0%;  Score 55;  DB 1;  Length 9;
  Best Local Similarity  100.0%;  Pred. No. 1.6e+05;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 YPYDVPDYA 9
          |||||||||
Db      1 YPYDVPDYA 9


RESULT   8
US-08-470-837-28
; Sequence 28, Application US/08470837
; Patent No. 5800811
;  GENERAL INFORMATION:
;    APPLICANT:  Nimni, Marcel E.
;    APPLICANT:  Hall, Frederick L.
;    APPLICANT:  Tuan, Tai-Lan
;    APPLICANT:  Wu, Lingtao
;    APPLICANT:  Cheung, David T.
;    TITLE OF INVENTION:  Transforming Growth Factor B Fusion
;    TITLE OF INVENTION:  and
;    TITLE OF INVENTION:  Their Use in Wound Healing
;    NUMBER OF SEQUENCES:  34
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Merchant & Gould
;      STREET:  11150 Santa Monica Boulevard, Suite 400
;      CITY:  Los Angeles
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  90025-3395
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/470,837
;      FILING DATE:
;      CLASSIFICATION:  424
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Sharp, Janice A.
;      REGISTRATION NUMBER:  34,051
;      REFERENCE/DOCKET NUMBER:  30630-1US01
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  310-445-1140
;      TELEFAX:  310-445-9031
;  INFORMATION FOR SEQ ID NO:  28:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  9 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-470-837-28


  Query Match           100.0%;  Score 55;  DB 1;  Length 9;
  Best Local Similarity  100.0%;  Pred. No. 1.6e+05;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 YPYDVPDYA 9
          |||||||||
Db      1 YPYDVPDYA 9


RESULT   9
US-08-478-386A-19
; Sequence 19, Application US/08478386A
; Patent No. 5830462
;  GENERAL INFORMATION:
;    APPLICANT:  Crabtree, Gerald R.
;    APPLICANT:  Schreiber, Stuart L.
;    APPLICANT:  Spencer, David M.
;    APPLICANT:  Wandless, Thomas J.
;    APPLICANT:  Belshaw, Peter
;    TITLE OF INVENTION:  REGULATED TRANSCRIPTION OF TARGETED
;    TITLE OF INVENTION:  GENES AND OTHER BIOLOGICAL EVENTS
;    NUMBER OF SEQUENCES:  81
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  ARIAD Pharmaceuticals, Inc.
;      STREET:  26 Landsdowne Street
;      CITY:  Cambridge
;      STATE:  Massachusetts
;      COUNTRY:  USA
;      ZIP:  02139
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC/DOS/MS/DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/478,386A
;      FILING DATE:  07/JUN/1995
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Figg, E. Anthony
;      REGISTRATION NUMBER:  27,195
;      REFERENCE/DOCKET NUMBER:  2054-114A
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202) 783-6040
;      TELEFAX:  (202) 783-6031
;  INFORMATION FOR SEQ ID NO:  19:
```

```
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  9 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-478-386A-19


  Query Match           100.0%;  Score 55;  DB 2;  Length 9;
  Best Local Similarity  100.0%;  Pred. No. 1.6e+05;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 YPYDVPDYA 9
          |||||||||
Db      1 YPYDVPDYA 9


RESULT  10
US-08-292-597-19
; Sequence 19, Application US/08292597
; Patent No. 5834266
;  GENERAL INFORMATION:
;    APPLICANT:  Gerald R. Crabtree
;    APPLICANT:  Schreiber, Stuart L.
;    APPLICANT:  Spencer, David M.
;    APPLICANT:  Wandless, Thomas J.
;    APPLICANT:  Belshaw, Peter
;    TITLE OF INVENTION:  Regulated Apoptosis
;    NUMBER OF SEQUENCES:  81
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  ARIAD Pharmaceuticals, Inc.
;      STREET:  26 Landsdowne Street
;      CITY:  Cambridge
;      STATE:  Massachusetts
;      COUNTRY:  USA
;      ZIP:  02139
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC/DOS/MS/DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/292,597
;      FILING DATE:  18/AUG/1994
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Figg, E. Anthony
;      REGISTRATION NUMBER:  27,195
;      REFERENCE/DOCKET NUMBER:  2054-108A
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202) 783-6040
;      TELEFAX:  (202) 783-6031
;  INFORMATION FOR SEQ ID NO:  19:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  9 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-292-597-19


  Query Match           100.0%;  Score 55;  DB 2;  Length 9;
  Best Local Similarity  100.0%;  Pred. No. 1.6e+05;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 YPYDVPDYA 9
          |||||||||
Db      1 YPYDVPDYA 9


RESULT  11
US-08-462-498-6
; Sequence 6, Application US/08462498
; Patent No. 5852169
;  GENERAL INFORMATION:
;    APPLICANT:  Roger Brent
;    APPLICANT:  Antonis S. Zervos
;    TITLE OF INVENTION:  MAX-INTERACTING PROTEINS AND RELATED
;    TITLE OF INVENTION:  MOLECULES AND METHODS
;    NUMBER OF SEQUENCES:  20
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Fish & Richardson
;      STREET:  225 Franklin Street
;      CITY:  Boston
;      STATE:  Massachusetts
;      COUNTRY:  U.S.A.
;      ZIP:  02110-2804
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5" Diskette, 1.44  Mb
;      COMPUTER:  IBM PS/2 Model 50Z or 55SX
;      OPERATING SYSTEM:  MS-DOS (Version 5.0)
;      SOFTWARE:  WordPerfect (Version 5.1)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/462,498
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/011,398
;      FILING DATE:  29 JAN 1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Paul T. Clark
;      REGISTRATION NUMBER:  30,162
;      REFERENCE/DOCKET NUMBER:  00786/160001
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617) 542-5070
;      TELEFAX:  (617) 542-8906
;      TELEX:  200154
;  INFORMATION FOR SEQ ID NO:  6:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  9
;      TYPE:  amino acid
;      STRANDEDNESS:  N/A
;      TOPOLOGY:  linear
US-08-462-498-6


  Query Match           100.0%;  Score 55;  DB 2;  Length 9;
  Best Local Similarity  100.0%;  Pred. No. 1.6e+05;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 YPYDVPDYA 9
          |||||||||
Db      1 YPYDVPDYA 9


RESULT  12
US-08-922-267A-60
; Sequence 60, Application US/08922267A
; Patent No. 5861239
;  GENERAL INFORMATION:
;    APPLICANT:  Kleyn, Patrick W.
;    APPLICANT:  Moore, Karen J.
;    TITLE OF INVENTION:  COMPOSITIONS FOR THE TREATMENT AND
;    TITLE OF INVENTION:  DIAGNOSIS OF BODY WEIGHT DISORDERS, INCLUDING OBESITY
;    NUMBER OF SEQUENCES:  82
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Pennie & Edmonds LLP
;      STREET:  1155 Avenue of the Americas
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  U.S.A.
;      ZIP:  10036-2711
;    COMPUTER READABLE FORM:
```

```
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.30
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/922,267A
;       FILING DATE:  2-SEP-1997
;       CLASSIFICATION:  530
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 08/829,553
;       FILING DATE:  28-MAR-1997
;       CLASSIFICATION:  530
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 08/631,200
;       FILING DATE:  12-APR-1996
;       CLASSIFICATION:  530
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Coruzzi, Laura A.
;       REGISTRATION NUMBER:  30,742
;       REFERENCE/DOCKET NUMBER:  7853-085
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (212) 790-9090
;       TELEFAX:  (212) 869-9741/8864
;       TELEX:  66141 PENNIE
;   INFORMATION FOR SEQ ID NO:  60:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  9 amino acids
;       TYPE:  amino acid
;       STRANDEDNESS:
;       TOPOLOGY:  unknown
US-08-922-267A-60


  Query Match           100.0%;  Score 55;  DB 2;  Length 9;
  Best Local Similarity  100.0%;  Pred. No. 1.6e+05;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 YPYDVPDYA 9
          |||||||||
Db      1 YPYDVPDYA 9


RESULT  13
US-08-388-653-19
; Sequence 19, Application US/08388653
; Patent No. 5869337
;  GENERAL INFORMATION:
;     APPLICANT:  Crabtree, Gerald R.
;     APPLICANT:  Schreiber, Stuart L.
;     APPLICANT:  Spencer, David M.
;     APPLICANT:  Wandless, Thomas J.
;     APPLICANT:  Belshaw, Peter
;     TITLE OF INVENTION:  REGULATED TRANSCRIPTION OF TARGETED
;     TITLE OF INVENTION:  GENES AND OTHER BIOLOGICAL EVENTS
;     NUMBER OF SEQUENCES:  81
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  ARIAD Pharmaceuticals, Inc.
;       STREET:  26 Landsdowne Street
;       CITY:  Cambridge
;       STATE:  Massachusetts
;       COUNTRY:  USA
;       ZIP:  02139
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC/DOS/MS/DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.25
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/388,653
;       FILING DATE:  14-FEB-1995
;       CLASSIFICATION:  514
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 08/478,386
;       FILING DATE:  07-JUN-1995
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Figg, E. Anthony
;       REGISTRATION NUMBER:  27,195
;       REFERENCE/DOCKET NUMBER:  2054-114A
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (202) 783-6040
;       TELEFAX:  (202) 783-6031
;   INFORMATION FOR SEQ ID NO:  19:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  9 amino acids
;       TYPE:  amino acid
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  protein
US-08-388-653-19


  Query Match           100.0%;  Score 55;  DB 2;  Length 9;
  Best Local Similarity  100.0%;  Pred. No. 1.6e+05;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 YPYDVPDYA 9
          |||||||||
Db      1 YPYDVPDYA 9


RESULT  14
US-08-473-985-19
; Sequence 19, Application US/08473985
; Patent No. 5871753
;  GENERAL INFORMATION:
;     APPLICANT:  Crabtree, Gerald R.
;     APPLICANT:  Schreiber, Stuart L.
;     APPLICANT:  Spencer, David M.
;     APPLICANT:  Wandless, Thomas J.
;     APPLICANT:  Belshaw, Peter
;     APPLICANT:  Ho, Steffan
;     TITLE OF INVENTION:  Regulated Transcription of Targeted Genes and
;     TITLE OF INVENTION:  Other Biological Events
;     NUMBER OF SEQUENCES:  66
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  Campbell and Flores
;       STREET:  4370 La Jolla Village Drive, Suite 700
;       CITY:  San Diego
;       STATE:  California
;       COUNTRY:  USA
;       ZIP:  92122
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.25
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/473,985
;       FILING DATE:
;       CLASSIFICATION:
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 08/179,748
;       FILING DATE:  07-JAN-1994
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Campbell, Cathryn A.
;       REGISTRATION NUMBER:  31,815
;       REFERENCE/DOCKET NUMBER:  P-SU 9863
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (619) 535-9001
;       TELEFAX:  (619) 535-8949
;   INFORMATION FOR SEQ ID NO:  19:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  9 amino acids
;       TYPE:  amino acid
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  protein
US-08-473-985-19
```

```
  Query Match           100.0%;  Score 55;  DB 2;  Length 9;
  Best Local Similarity  100.0%;  Pred. No. 1.6e+05;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 YPYDVPDYA 9
          |||||||||
Db      1 YPYDVPDYA 9


RESULT  15
US-08-690-011A-17
; Sequence 17, Application US/08690011A
; Patent No. 5942433
;  GENERAL INFORMATION:
;    APPLICANT:  VINSON, Charles R.
;    APPLICANT:  KRYLOV, Dmitry
;    TITLE OF INVENTION:  EXTENSION OF A PROTEIN-PROTEIN
;    TITLE OF INVENTION:  INTERACTION SURFACE TO INACTIVATE THE FUNCTION OF A
;    TITLE OF INVENTION:  CELLULAR PROTEIN
;    NUMBER OF SEQUENCES:  60
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  MORGAN & FINNEGAN, L.L.P.
;      STREET:  345 Park Avenue
;      CITY:  New York
;      STATE:  NY
;      COUNTRY:  USA
;      ZIP:  10154-0053
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30B
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/690,011A
;      FILING DATE:  31-JUL-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  60/001,654
;      FILING DATE:  31-JUL-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  60/018,496
;      FILING DATE:  29-MAY-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Serunian, Leslie A.
;      REGISTRATION NUMBER:  35,353
;      REFERENCE/DOCKET NUMBER:  2026-4199US1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212)758-4800
;      TELEFAX:  (212)751-6849
;  INFORMATION FOR SEQ ID NO:  17:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  9 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  unknown
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-690-011A-17


  Query Match           100.0%;  Score 55;  DB 2;  Length 9;
  Best Local Similarity  100.0%;  Pred. No. 1.6e+05;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 YPYDVPDYA 9
          |||||||||
Db      1 YPYDVPDYA 9


Search completed: April 11, 2002, 10:14:43
Job time: 33 sec
```

GenCore version 4.5

OM protein - protein search, using sw model

Run on: April 11, 2002, 10:14:10 ; Search time 27.06 Seconds
(without alignments)
25.335 Million cell updates/sec

Title: US-09-284-787-1
Perfect score: 55
Sequence: 1 YPYDVPDYA 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_68:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	55	100.0	330	2	JQ2374	hemagglutinin - in
2	55	100.0	330	2	JQ2375	hemagglutinin - in
3	55	100.0	331	2	JQ2377	hemagglutinin - in
4	55	100.0	331	2	JQ2378	hemagglutinin - in
5	55	100.0	347	2	S52173	hemagglutinin - in
6	55	100.0	347	2	S52174	hemagglutinin - in
7	55	100.0	347	2	S52175	hemagglutinin - in
8	55	100.0	347	2	S52176	hemagglutinin - in
9	55	100.0	347	2	S52178	hemagglutinin - in
10	55	100.0	347	2	S52179	hemagglutinin - in
11	55	100.0	347	2	S52180	hemagglutinin - in
12	55	100.0	347	2	S52182	hemagglutinin - in
13	55	100.0	347	2	S52183	hemagglutinin - in
14	55	100.0	347	2	S52185	hemagglutinin - in
15	55	100.0	347	2	S52186	hemagglutinin - in
16	55	100.0	347	2	S52187	hemagglutinin - in
17	55	100.0	347	2	S52188	hemagglutinin - in
18	55	100.0	347	2	S52189	hemagglutinin - in
19	55	100.0	347	2	S52191	hemagglutinin - in
20	55	100.0	347	2	S52195	hemagglutinin - in
21	55	100.0	347	2	S52197	hemagglutinin - in
22	55	100.0	347	2	S52199	hemagglutinin - in
23	55	100.0	362	2	S38637	hemagglutinin - in
24	55	100.0	550	1	HMIVS2	hemagglutinin prec
25	55	100.0	550	1	HMIVS3	hemagglutinin prec
26	55	100.0	550	1	HMIV77	hemagglutinin prec
27	55	100.0	550	1	HMIV80	hemagglutinin prec
28	55	100.0	550	1	HMIV33	hemagglutinin prec
29	55	100.0	550	1	HMIV89	hemagglutinin prec
30	55	100.0	550	1	HMIV98	hemagglutinin prec
31	55	100.0	550	1	HMIV15	hemagglutinin prec
32	55	100.0	550	1	HMIV86	hemagglutinin prec
33	55	100.0	550	2	JQ1153	hemagglutinin prec
34	55	100.0	550	2	JQ1154	hemagglutinin prec
35	55	100.0	550	2	JQ1155	hemagglutinin prec
36	55	100.0	550	2	JQ1156	hemagglutinin prec
37	55	100.0	565	1	HMIVE2	hemagglutinin prec
38	55	100.0	565	1	HMIVE4	hemagglutinin prec
39	55	100.0	566	1	HMIVH	hemagglutinin prec
40	55	100.0	566	1	HMIVHA	hemagglutinin prec
41	55	100.0	566	1	HMIVHM	hemagglutinin prec
42	55	100.0	566	1	HMIV6	hemagglutinin prec
43	55	100.0	567	1	HMIVV	hemagglutinin prec
44	54	98.2	565	1	HMIVE1	hemagglutinin prec
45	54	98.2	565	1	HMIVE5	hemagglutinin prec

ALIGNMENTS

RESULT 1
JQ2374
hemagglutinin - influenza A virus (strain SN1289)
C;Species: influenza A virus
C;Date: 14-Jul-1994 #sequence_revision 14-Jul-1994 #text_change 07-May-1999
C;Accession: JQ2374
R;Rocha, E.P.; Xu, X.; Hall, H.E.; Allen, J.R.; Regnery, H.L.; Cox, N.J.
J. Gen. Virol. 74, 2513-2518, 1993
A;Title: Comparison of 10 influenza A (H1N1 and H3N2) haemagglutinin sequences obtain
A;Reference number: JQ2369; MUID:94065682
A;Accession: JQ2374
A;Molecule type: mRNA
A;Residues: 1-330 <ROC>
A;Cross-references: GB:L20115
A;Experimental source: subtype H3N2
C;Superfamily: influenza virus hemagglutinin

Query Match 100.0%; Score 55; DB 2; Length 330;
Best Local Similarity 100.0%; Pred. No. 0.022;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 YPYDVPDYA 9
||||||||| |
Db 98 YPYDVPDYA 106

RESULT 2
JQ2375
hemagglutinin - influenza A virus (strain SN1389 and IN0190)
C;Species: influenza A virus
C;Date: 14-Jul-1994 #sequence_revision 14-Jul-1994 #text_change 07-May-1999
C;Accession: JQ2375; JQ2376
R;Rocha, E.P.; Xu, X.; Hall, H.E.; Allen, J.R.; Regnery, H.L.; Cox, N.J.
J. Gen. Virol. 74, 2513-2518, 1993
A;Title: Comparison of 10 influenza A (H1N1 and H3N2) haemagglutinin sequences obtain
A;Reference number: JQ2369; MUID:94065682
A;Accession: JQ2375
A;Molecule type: mRNA
A;Residues: 1-330 <ROC>
A;Cross-references: GB:L20119
A;Experimental source: subtype H3N2
C;Superfamily: influenza virus hemagglutinin

Query Match 100.0%; Score 55; DB 2; Length 330;
Best Local Similarity 100.0%; Pred. No. 0.022;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 YPYDVPDYA 9
||||||||| |
Db 98 YPYDVPDYA 106

```
RESULT   3
JQ2377
hemagglutinin - influenza A virus (strain SD0191)
C;Species: influenza A virus
C;Date: 14-Jul-1994 #sequence_revision 14-Jul-1994 #text_change 07-May-1999
C;Accession: JQ2377
R;Rocha, E.P.; Xu, X.; Hall, H.E.; Allen, J.R.; Regnery, H.L.; Cox, N.J.
J. Gen. Virol. 74, 2513-2518, 1993
A;Title: Comparison of 10 influenza A (H1N1 and H3N2) haemagglutinin sequences obtained
A;Reference number: JQ2369; MUID:94065682
A;Accession: JQ2377
A;Molecule type: mRNA
A;Residues: 1-331 <ROC>
A;Cross-references: GB:L20105
A;Experimental source: subtype H3N2
C;Superfamily: influenza virus hemagglutinin


  Query Match            100.0%;  Score 55;  DB 2;  Length 331;
  Best Local Similarity  100.0%;  Pred. No. 0.022;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 YPYDVPDYA 9
            |||||||||
Db      98 YPYDVPDYA 106


RESULT   4
JQ2378
hemagglutinin - influenza A virus (strain HI0191)
C;Species: influenza A virus
C;Date: 14-Jul-1994 #sequence_revision 14-Jul-1994 #text_change 07-May-1999
C;Accession: JQ2378
R;Rocha, E.P.; Xu, X.; Hall, H.E.; Allen, J.R.; Regnery, H.L.; Cox, N.J.
J. Gen. Virol. 74, 2513-2518, 1993
A;Title: Comparison of 10 influenza A (H1N1 and H3N2) haemagglutinin sequences obtained
A;Reference number: JQ2369; MUID:94065682
A;Accession: JQ2378
A;Molecule type: mRNA
A;Residues: 1-331 <ROC>
A;Cross-references: GB:L20102
A;Experimental source: subtype H3N2
C;Superfamily: influenza virus hemagglutinin


  Query Match            100.0%;  Score 55;  DB 2;  Length 331;
  Best Local Similarity  100.0%;  Pred. No. 0.022;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 YPYDVPDYA 9
            |||||||||
Db      98 YPYDVPDYA 106


RESULT   5
S52173
hemagglutinin - influenza A virus (fragment)
C;Species: influenza A virus
C;Date: 07-May-1995 #sequence_revision 23-Feb-1996 #text_change 23-Mar-2001
C;Accession: S52173; S52190; S52192; S52177; S52181; S52194; S52196; S52184; S52193
R;Ellis, J.S.
submitted to the EMBL Data Library, October 1994
A;Description: Antigenic and genetic variation in the haemagglutinin of recently circula
A;Reference number: S52173
A;Accession: S52173
A;Molecule type: mRNA
A;Residues: 1-347 <ELL>
A;Cross-references: EMBL:Z46392; NID:g609020; PIDN:CAA86527.1; PID:g609021
A;Experimental source: isolate A/Beijing/32/92
A;Accession: S52190
A;Molecule type: mRNA
A;Residues: 1-347 <EL2>
A;Cross-references: EMBL:Z46410; NID:g609053; PIDN:CAA86545.1; PID:g609054
A;Experimental source: isolate A/Hong Kong/23/92
A;Accession: S52192
A;Molecule type: mRNA
A;Residues: 1-347 <EL3>
A;Cross-references: EMBL:Z46409; NID:g609057; PIDN:CAA86544.1; PID:g609058
A;Experimental source: isolate A/Hong Kong/34/90
A;Accession: S52177
A;Molecule type: mRNA
A;Residues: 1-134,'K',136-156,'L',158-173,'F',175-188,'S',190-218,'F',220-225,'L',227
A;Cross-references: EMBL:Z46395; NID:g609027; PIDN:CAA86530.1; PID:g1228087
A;Experimental source: isolate A/England/269/93
A;Accession: S52181
A;Molecule type: mRNA
A;Residues: 1-134,'K',136-156,'L',158-173,'F',175-188,'S',190-218,'F',220-225,'L',227
A;Cross-references: EMBL:Z46399; NID:g609035; PIDN:CAA86534.1; PID:g609036
A;Experimental source: isolate A/England/328/93
A;Accession: S52194
A;Molecule type: mRNA
A;Residues: 1-134,'K',136-156,'L',158-173,'F',175-188,'S',190-218,'F',220-225,'L',227
A;Cross-references: EMBL:Z46414; NID:g609061; PIDN:CAA86549.1; PID:g609062
A;Experimental source: isolate A/Scotland/160/93
A;Accession: S52196
A;Molecule type: mRNA
A;Residues: 1-134,'K',136-156,'L',158-173,'F',175-188,'S',190-218,'F',220-225,'L',227
A;Cross-references: EMBL:Z46415; NID:g609065; PIDN:CAA86550.1; PID:g609066
A;Experimental source: isolate A/Scotland/174/93
A;Accession: S52184
A;Molecule type: mRNA
A;Residues: 1-74,'N',76-120,'T',122-156,'L',158-173,'F',175-188,'S',190-200,'K',202-2
A;Cross-references: EMBL:Z46402; NID:g609041; PIDN:CAA86537.1; PID:g609042
A;Experimental source: isolate A/England/471/93
A;Accession: S52193
A;Molecule type: mRNA
A;Residues: 1-74,'N',76-120,'T',122-156,'L',158-173,'F',175-188,'S',190-200,'K',202-2
A;Cross-references: EMBL:Z46413; NID:g609059; PIDN:CAA86548.1; PID:g609060
A;Experimental source: isolate A/Scotland/142/93
C;Superfamily: influenza virus hemagglutinin


  Query Match            100.0%;  Score 55;  DB 2;  Length 347;
  Best Local Similarity  100.0%;  Pred. No. 0.023;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 YPYDVPDYA 9
            |||||||||
Db      98 YPYDVPDYA 106


RESULT   6
S52174
hemagglutinin - influenza A virus (isolate A/Beijing/353/89) (fragment)
C;Species: influenza A virus
A;Variety: isolate A/Beijing/353/89
C;Date: 07-May-1995 #sequence_revision 23-Feb-1996 #text_change 20-Sep-1999
C;Accession: S52174
R;Ellis, J.S.
submitted to the EMBL Data Library, October 1994
A;Description: Antigenic and genetic variation in the haemagglutinin of recently circ
A;Reference number: S52173
A;Accession: S52174
A;Molecule type: mRNA
A;Residues: 1-347 <ELL>
A;Cross-references: EMBL:Z46391; NID:g609022; PIDN:CAA86526.1; PID:g609023
A;Experimental source: isolate A/Beijing/353/89
C;Superfamily: influenza virus hemagglutinin


  Query Match            100.0%;  Score 55;  DB 2;  Length 347;
  Best Local Similarity  100.0%;  Pred. No. 0.023;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;
```

```
Qy      1 YPYDVPDYA 9
          |||||||||
Db     98 YPYDVPDYA 106


RESULT   7
S52175
hemagglutinin - influenza A virus (isolate A/England/1/93) (fragment)
C;Species: influenza A virus
A;Variety: isolate A/England/1/93
C;Date: 07-May-1995 #sequence_revision 23-Feb-1996 #text_change 23-Mar-2001
C;Accession: S52175
R;Ellis, J.S.
submitted to the EMBL Data Library, October 1994
A;Description: Antigenic and genetic variation in the haemagglutinin of recently circula
A;Reference number: S52173
A;Accession: S52175
A;Molecule type: mRNA
A;Residues: 1-347 <ELL>
A;Cross-references: EMBL:Z46393; NID:g609024; PIDN:CAA86528.1; PID:g940547
A;Experimental source: isolate A/England/1/93
C;Superfamily: influenza virus hemagglutinin


  Query Match             100.0%;  Score 55;  DB 2;  Length 347;
  Best Local Similarity   100.0%;  Pred. No. 0.023;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 YPYDVPDYA 9
          |||||||||
Db     98 YPYDVPDYA 106


RESULT   8
S52176
hemagglutinin - influenza A virus (isolate A/England/247/93) (fragment)
C;Species: influenza A virus
A;Variety: isolate A/England/247/93
C;Date: 07-May-1995 #sequence_revision 23-Feb-1996 #text_change 23-Mar-2001
C;Accession: S52176
R;Ellis, J.S.
submitted to the EMBL Data Library, October 1994
A;Description: Antigenic and genetic variation in the haemagglutinin of recently circula
A;Reference number: S52173
A;Accession: S52176
A;Molecule type: mRNA
A;Residues: 1-347 <ELL>
A;Cross-references: EMBL:Z46394; NID:g609025; PIDN:CAA86529.1; PID:g609026
A;Experimental source: isolate A/England/247/93
C;Superfamily: influenza virus hemagglutinin


  Query Match             100.0%;  Score 55;  DB 2;  Length 347;
  Best Local Similarity   100.0%;  Pred. No. 0.023;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 YPYDVPDYA 9
          |||||||||
Db     98 YPYDVPDYA 106


RESULT   9
S52178
hemagglutinin - influenza A virus (isolate A/England/284/93) (fragment)
C;Species: influenza A virus
A;Variety: isolate A/England/284/93
C;Date: 07-May-1995 #sequence_revision 23-Feb-1996 #text_change 23-Mar-2001
C;Accession: S52178
R;Ellis, J.S.
submitted to the EMBL Data Library, October 1994
A;Description: Antigenic and genetic variation in the haemagglutinin of recently circula
A;Reference number: S52173
A;Accession: S52178
A;Molecule type: mRNA
A;Residues: 1-347 <ELL>
A;Cross-references: EMBL:Z46396; NID:g609029; PIDN:CAA86531.1; PID:g609030
A;Experimental source: isolate A/England/284/93
C;Superfamily: influenza virus hemagglutinin


  Query Match             100.0%;  Score 55;  DB 2;  Length 347;
  Best Local Similarity   100.0%;  Pred. No. 0.023;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 YPYDVPDYA 9
          |||||||||
Db     98 YPYDVPDYA 106


RESULT  10
S52179
hemagglutinin - influenza A virus (isolate A/England/286/93) (fragment)
C;Species: influenza A virus
A;Variety: isolate A/England/286/93
C;Date: 07-May-1995 #sequence_revision 23-Feb-1996 #text_change 23-Mar-2001
C;Accession: S52179
R;Ellis, J.S.
submitted to the EMBL Data Library, October 1994
A;Description: Antigenic and genetic variation in the haemagglutinin of recently circ
A;Reference number: S52173
A;Accession: S52179
A;Molecule type: mRNA
A;Residues: 1-347 <ELL>
A;Cross-references: EMBL:Z46397; NID:g609031; PIDN:CAA86532.1; PID:g609032
A;Experimental source: isolate A/England/286/93
C;Superfamily: influenza virus hemagglutinin


  Query Match             100.0%;  Score 55;  DB 2;  Length 347;
  Best Local Similarity   100.0%;  Pred. No. 0.023;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 YPYDVPDYA 9
          |||||||||
Db     98 YPYDVPDYA 106


RESULT  11
S52180
hemagglutinin - influenza A virus (isolate A/England/289/93) (fragment)
C;Species: influenza A virus
A;Variety: isolate A/England/289/93
C;Date: 07-May-1995 #sequence_revision 23-Feb-1996 #text_change 23-Mar-2001
C;Accession: S52180
R;Ellis, J.S.
submitted to the EMBL Data Library, October 1994
A;Description: Antigenic and genetic variation in the haemagglutinin of recently circ
A;Reference number: S52173
A;Accession: S52180
A;Molecule type: mRNA
A;Residues: 1-347 <ELL>
A;Cross-references: EMBL:Z46398; NID:g609033; PIDN:CAA86533.1; PID:g609034
A;Experimental source: isolate A/England/289/93
C;Superfamily: influenza virus hemagglutinin


  Query Match             100.0%;  Score 55;  DB 2;  Length 347;
  Best Local Similarity   100.0%;  Pred. No. 0.023;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 YPYDVPDYA 9
          |||||||||
Db     98 YPYDVPDYA 106


RESULT  12
```

S52182
hemagglutinin - influenza A virus (isolate A/England/328/93) (fragment)
C;Species: influenza A virus
A;Variety: isolate A/England/328/93
C;Date: 07-May-1995 #sequence_revision 23-Feb-1996 #text_change 23-Mar-2001
C;Accession: S52182
R;Ellis, J.S.
submitted to the EMBL Data Library, October 1994
A;Description: Antigenic and genetic variation in the haemagglutinin of recently circula
A;Reference number: S52173
A;Accession: S52182
A;Molecule type: mRNA
A;Residues: 1-347 <ELL>
A;Cross-references: EMBL:Z46400; NID:g609037; PIDN:CAA86535.1; PID:g609038
A;Experimental source: isolate A/England/346/93
C;Superfamily: influenza virus hemagglutinin

Query Match 100.0%; Score 55; DB 2; Length 347;
Best Local Similarity 100.0%; Pred. No. 0.023;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 YPYDVPDYA 9
|||||||||
Db 98 YPYDVPDYA 106

RESULT 13
S52183
hemagglutinin - influenza A virus (isolate A/England/347/93) (fragment)
C;Species: influenza A virus
A;Variety: isolate A/England/347/93
C;Date: 07-May-1995 #sequence_revision 23-Feb-1996 #text_change 23-Mar-2001
C;Accession: S52183
R;Ellis, J.S.
submitted to the EMBL Data Library, October 1994
A;Description: Antigenic and genetic variation in the haemagglutinin of recently circula
A;Reference number: S52173
A;Accession: S52183
A;Molecule type: mRNA
A;Residues: 1-347 <ELL>
A;Cross-references: EMBL:Z46401; NID:g609039; PIDN:CAA86536.1; PID:g609040
A;Experimental source: isolate A/England/347/93
C;Superfamily: influenza virus hemagglutinin

Query Match 100.0%; Score 55; DB 2; Length 347;
Best Local Similarity 100.0%; Pred. No. 0.023;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 YPYDVPDYA 9
|||||||||
Db 98 YPYDVPDYA 106

RESULT 14
S52185
hemagglutinin - influenza A virus (isolate A/England/67/94) (fragment)
C;Species: influenza A virus
A;Variety: isolate A/England/67/94
C;Date: 07-May-1995 #sequence_revision 23-Feb-1996 #text_change 23-Mar-2001
C;Accession: S52185
R;Ellis, J.S.
submitted to the EMBL Data Library, October 1994
A;Description: Antigenic and genetic variation in the haemagglutinin of recently circula
A;Reference number: S52173
A;Accession: S52185
A;Molecule type: mRNA
A;Residues: 1-347 <ELL>
A;Cross-references: EMBL:Z46403; NID:g609043; PIDN:CAA86538.1; PID:g609044
A;Experimental source: isolate A/England/67/94
C;Superfamily: influenza virus hemagglutinin

Query Match 100.0%; Score 55; DB 2; Length 347;
Best Local Similarity 100.0%; Pred. No. 0.023;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 YPYDVPDYA 9
|||||||||
Db 98 YPYDVPDYA 106

RESULT 15
S52186
hemagglutinin - influenza A virus (isolate A/England/68/94) (fragment)
C;Species: influenza A virus
A;Variety: isolate A/England/68/94
C;Date: 07-May-1995 #sequence_revision 23-Feb-1996 #text_change 23-Mar-2001
C;Accession: S52186
R;Ellis, J.S.
submitted to the EMBL Data Library, October 1994
A;Description: Antigenic and genetic variation in the haemagglutinin of recently circ
A;Reference number: S52173
A;Accession: S52186
A;Molecule type: mRNA
A;Residues: 1-347 <ELL>
A;Cross-references: EMBL:Z46404; NID:g609045; PIDN:CAA86539.1; PID:g609046
A;Experimental source: isolate A/England/68/94
C;Superfamily: influenza virus hemagglutinin

Query Match 100.0%; Score 55; DB 2; Length 347;
Best Local Similarity 100.0%; Pred. No. 0.023;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 YPYDVPDYA 9
|||||||||
Db 98 YPYDVPDYA 106

Search completed: April 11, 2002, 10:16:16
Job time: 126 sec

GenCore version 4.5

OM protein - protein search, using sw model

Run on: April 11, 2002, 10:14:50 ; Search time 18.1 Seconds
(without alignments)
18.231 Million cell updates/sec

Title: US-09-284-787-1
Perfect score: 55
Sequence: 1 YPYDVPDYA 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 100059 seqs, 36664827 residues

Total number of hits satisfying chosen parameters: 100059

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID		Description
1	55	100.0	249	1	HEMA_IAX31	P03438	influenza a
2	55	100.0	328	1	HEMA_IAEN6	P04664	influenza a
3	55	100.0	328	1	HEMA_IAQU7	P04663	influenza a
4	55	100.0	550	1	HEMA_IABAN	P03441	influenza a
5	55	100.0	550	1	HEMA_IADH1	P12582	influenza a
6	55	100.0	550	1	HEMA_IADH2	P12583	influenza a
7	55	100.0	550	1	HEMA_IADH3	P12584	influenza a
8	55	100.0	550	1	HEMA_IADH4	P12585	influenza a
9	55	100.0	550	1	HEMA_IADH6	P12587	influenza a
10	55	100.0	550	1	HEMA_IADH7	P12588	influenza a
11	55	100.0	550	1	HEMA_IADHK	P43257	influenza a
12	55	100.0	550	1	HEMA_IADHL	P43258	influenza a
13	55	100.0	550	1	HEMA_IADHM	P43259	influenza a
14	55	100.0	550	1	HEMA_IAGHK	P43260	influenza a
15	55	100.0	550	1	HEMA_IAME6	P12589	influenza a
16	55	100.0	550	1	HEMA_IAZH2	P11133	influenza a
17	55	100.0	550	1	HEMA_IAZH3	P11134	influenza a
18	55	100.0	565	1	HEMA_IAHAL	P16994	influenza a
19	55	100.0	565	1	HEMA_IAHMI	P15658	influenza a
20	55	100.0	566	1	HEMA_IAAIC	P03437	influenza a
21	55	100.0	566	1	HEMA_IADM2	P26135	influenza a
22	55	100.0	566	1	HEMA_IAEN7	P03440	influenza a
23	55	100.0	566	1	HEMA_IAMAO	P26138	influenza a
24	55	100.0	566	1	HEMA_IAME1	P03449	influenza a
25	55	100.0	566	1	HEMA_IAME2	P03439	influenza a
26	55	100.0	566	1	HEMA_IANT6	P03436	influenza a
27	55	100.0	566	1	HEMA_IAUDO	P19106	influenza a
28	55	100.0	566	1	HEMA_IAZCO	P26139	influenza a
29	55	100.0	567	1	HEMA_IAVI7	P03435	influenza a
30	54	98.2	565	1	HEMA_IAHFO	P16995	influenza a
31	54	98.2	565	1	HEMA_IAHK6	P19699	influenza a
32	54	98.2	565	1	HEMA_IAHK7	P16996	influenza a
33	54	98.2	565	1	HEMA_IAHNM	P16997	influenza a
34	54	98.2	565	1	HEMA_IAHRO	P16998	influenza a
35	54	98.2	565	1	HEMA_IAHSA	P16999	influenza a
36	54	98.2	565	1	HEMA_IAHSU	Q08011	influenza a
37	54	98.2	565	1	HEMA_IAHTE	P17001	influenza a
38	54	98.2	565	1	HEMA_IAHUR	P17002	influenza a
39	54	98.2	566	1	HEMA_IADU3	P03442	influenza a
40	51	92.7	566	1	HEMA_IADA3	P26134	influenza a
41	51	92.7	566	1	HEMA_IAZUK	P26141	influenza a
42	50	90.9	565	1	HEMA_IAHTO	P17000	influenza a
43	47	85.5	550	1	HEMA_IADH5	P12586	influenza a
44	47	85.5	563	1	HEMA_IAMAA	P26137	influenza a
45	47	85.5	564	1	HEMA_IABUD	P19694	influenza a

ALIGNMENTS

RESULT 1
HEMA_IAX31
ID HEMA_IAX31 STANDARD; PRT; 249 AA.
AC P03438;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE HEMAGGLUTININ PRECURSOR [CONTAINS: HEMAGGLUTININ HA1 CHAIN]
DE (FRAGMENT).
GN HA.
OS Influenza A virus (strain X-31).
OC Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;
OC Influenza virus A and B group; Influenza A viruses;
OC Influenza A virus.
OX NCBI_TaxID=11489;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=81030852; PubMed=7421990;
RA Gething M.-J., Bye J., Skehel J.J., Waterfield M.;
RT "Cloning and DNA sequence of double-stranded copies of haemagglutinin
RT genes from H2 and H3 strains elucidates antigenic shift and drift in
RT human influenza virus.";
RL Nature 287:301-306(1980).
CC -!- FUNCTION: HEMAGGLUTININ IS RESPONSIBLE FOR ATTACHING THE VIRUS TO
CC CELL RECEPTORS AND FOR INITIATING INFECTION.
CC -!- SUBUNIT: HOMOTRIMER. EACH OF THE MONOMERS IS FORMED BY TWO CHAINS
CC (HA1 AND HA2) LINKED BY A DISULFIDE BOND.
CC -!- SIMILARITY: BELONGS TO THE INFLUENZA HEMAGGLUTININ FAMILY.
DR PIR; A04051; HMIVH.
DR PIR; A93233; A93233.
DR HSSP; P03437; 5HMG.
DR InterPro; IPR001364; Hemagglutn.
DR Pfam; PF00509; Hemagglutinin; 1.
DR ProDom; PD000225; Hemagglutn; 1.
KW Envelope protein; Hemagglutinin; Glycoprotein; Signal.
FT SIGNAL 1 16
FT CHAIN 17 >249 HEMAGGLUTININ HA1 CHAIN.
FT CARBOHYD 24 24 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 38 38 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 54 54 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 97 97 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 181 181 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT NON_TER 249 249
SQ SEQUENCE 249 AA; 27373 MW; 9537AA970BB79183 CRC64;

Query Match 100.0%; Score 55; DB 1; Length 249;
Best Local Similarity 100.0%; Pred. No. 0.01;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 YPYDVPDYA 9
|||||||||
Db 114 YPYDVPDYA 122

RESULT 2

```
HEMA_IAEN6
ID   HEMA_IAEN6     STANDARD;      PRT;   328 AA.
AC   P04664;
DT   13-AUG-1987 (Rel. 05, Created)
DT   13-AUG-1987 (Rel. 05, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   HEMAGGLUTININ [CONTAINS: HEMAGGLUTININ HA1 CHAIN] (FRAGMENT).
GN   HA.
OS   Influenza A virus (strain A/England/878/69).
OC   Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;
OC   Influenza virus A and B group; Influenza A viruses;
OC   Influenza A virus.
OX   NCBI_TaxID=11377;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=81194918; PubMed=6164798;
RA   Sleigh M.J., Both G.W., Underwood P.A., Bender V.J.;
RT   "Antigenic drift in the hemagglutinin of the Hong Kong influenza
RT   subtype: correlation of amino acid changes with alterations in viral
RT   antigenicity.";
RL   J. Virol. 37:845-853(1981).
CC   -!- FUNCTION: HEMAGGLUTININ IS RESPONSIBLE FOR ATTACHING THE VIRUS TO
CC       CELL RECEPTORS AND FOR INITIATING INFECTION.
CC   -!- SUBUNIT: HOMOTRIMER. EACH OF THE MONOMERS IS FORMED BY TWO CHAINS
CC       (HA1 AND HA2) LINKED BY A DISULFIDE BOND.
CC   -!- SIMILARITY: BELONGS TO THE INFLUENZA HEMAGGLUTININ FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; K03335; AAA43184.1; -.
DR   InterPro; IPR001364; Hemagglutn.
DR   Pfam; PF00509; Hemagglutinin; 1.
DR   ProDom; PD000225; Hemagglutn; 1.
KW   Envelope protein; Hemagglutinin; Glycoprotein.
FT   NON_TER       1      1
FT   CHAIN         1    328       HEMAGGLUTININ HA1 CHAIN.
FT   CARBOHYD      8      8       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     22     22       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     38     38       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     63     63       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    165    165       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    285    285       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   NON_TER     328    328
SQ   SEQUENCE   328 AA;  36072 MW;  9C3A86B3A8D856FE CRC64;


  Query Match             100.0%;  Score 55;  DB 1;  Length 328;
  Best Local Similarity   100.0%;  Pred. No. 0.013;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 YPYDVPDYA 9
           |||||||||
Db      98 YPYDVPDYA 106


RESULT   3
HEMA_IAQU7
ID   HEMA_IAQU7     STANDARD;      PRT;   328 AA.
AC   P04663;
DT   13-AUG-1987 (Rel. 05, Created)
DT   13-AUG-1987 (Rel. 05, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   HEMAGGLUTININ [CONTAINS: HEMAGGLUTININ HA1 CHAIN] (FRAGMENT).
GN   HA.
OS   Influenza A virus (strain A/Qu/7/70).
OC   Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;
OC   Influenza virus A and B group; Influenza A viruses;
OC   Influenza A virus.
OX   NCBI_TaxID=11456;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=81194918; PubMed=6164798;
RA   Sleigh M.J., Both G.W., Underwood P.A., Bender V.J.;
RT   "Antigenic drift in the hemagglutinin of the Hong Kong influenza
RT   subtype: correlation of amino acid changes with alterations in viral
RT   antigenicity.";
RL   J. Virol. 37:845-853(1981).
CC   -!- FUNCTION: HEMAGGLUTININ IS RESPONSIBLE FOR ATTACHING THE VIRUS TO
CC       CELL RECEPTORS AND FOR INITIATING INFECTION.
CC   -!- SUBUNIT: HOMOTRIMER. EACH OF THE MONOMERS IS FORMED BY TWO CHAINS
CC       (HA1 AND HA2) LINKED BY A DISULFIDE BOND.
CC   -!- SIMILARITY: BELONGS TO THE INFLUENZA HEMAGGLUTININ FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; K03338; AAA43195.1; -.
DR   InterPro; IPR001364; Hemagglutn.
DR   Pfam; PF00509; Hemagglutinin; 1.
DR   ProDom; PD000225; Hemagglutn; 1.
KW   Envelope protein; Hemagglutinin; Glycoprotein.
FT   NON_TER       1      1
FT   CHAIN         1    328       HEMAGGLUTININ HA1 CHAIN.
FT   CARBOHYD      8      8       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     22     22       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     38     38       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     81     81       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    165    165       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    285    285       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   NON_TER     328    328
SQ   SEQUENCE   328 AA;  36059 MW;  E30A962FD6DD805B CRC64;


  Query Match             100.0%;  Score 55;  DB 1;  Length 328;
  Best Local Similarity   100.0%;  Pred. No. 0.013;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 YPYDVPDYA 9
           |||||||||
Db      98 YPYDVPDYA 106


RESULT   4
HEMA_IABAN
ID   HEMA_IABAN     STANDARD;      PRT;   550 AA.
AC   P03441; Q83961; Q83962;
DT   21-JUL-1986 (Rel. 01, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   HEMAGGLUTININ PRECURSOR [CONTAINS: HEMAGGLUTININ HA1 CHAIN;
DE   HEMAGGLUTININ HA2 CHAIN] (FRAGMENT).
GN   HA.
OS   Influenza A virus (strain A/Bangkok/1/79).
OC   Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;
OC   Influenza virus A and B group; Influenza A viruses;
OC   Influenza A virus.
OX   NCBI_TaxID=11325;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=82033259; PubMed=6169840;
RA   Both G.W., Sleigh M.J.;
RT   "Conservation and variation in the hemagglutinins of Hong Kong
RT   subtype influenza viruses during antigenic drift.";
RL   J. Virol. 39:663-672(1981).
CC   -!- FUNCTION: HEMAGGLUTININ IS RESPONSIBLE FOR ATTACHING THE VIRUS TO
```

```
CC       CELL RECEPTORS AND FOR INITIATING INFECTION.
CC   -!- SUBUNIT: HOMOTRIMER. EACH OF THE MONOMERS IS FORMED BY TWO CHAINS
CC       (HA1 AND HA2) LINKED BY A DISULFIDE BOND.
CC   -!- SIMILARITY: BELONGS TO THE INFLUENZA HEMAGGLUTININ FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; J02092; AAA43182.1; -.
DR   PIR; A04052; HMIV6.
DR   InterPro; IPR001364; Hemagglutn.
DR   Pfam; PF00509; Hemagglutinin; 1.
DR   ProDom; PD000225; Hemagglutn; 1.
KW   Envelope protein; Hemagglutinin; Glycoprotein.
FT   NON_TER       1      1
FT   CHAIN         1    328       HEMAGGLUTININ HA1 CHAIN.
FT   CHAIN       330    550       HEMAGGLUTININ HA2 CHAIN.
FT   CARBOHYD      8      8       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     22     22       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     38     38       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     63     63       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    126    126       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    165    165       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    285    285       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    483    483       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   550 AA;  61699 MW;  FBD2EC200689CBE5 CRC64;


  Query Match             100.0%;  Score 55;  DB 1;  Length 550;
  Best Local Similarity   100.0%;  Pred. No. 0.023;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 YPYDVPDYA 9
           |||||||||
Db      98 YPYDVPDYA 106


RESULT   5
HEMA_IADH1
ID   HEMA_IADH1     STANDARD;      PRT;   550 AA.
AC   P12582; Q84021; Q84022;
DT   01-OCT-1989 (Rel. 12, Created)
DT   01-OCT-1989 (Rel. 12, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   HEMAGGLUTININ PRECURSOR [CONTAINS: HEMAGGLUTININ HA1 CHAIN;
DE   HEMAGGLUTININ HA2 CHAIN] (FRAGMENT).
GN   HA.
OS   Influenza A virus (strain A/Duck/Hokkaido/5/77).
OC   Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;
OC   Influenza virus A and B group; Influenza A viruses;
OC   Influenza A virus.
OX   NCBI_TaxID=11357;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=87265458; PubMed=2440178;
RA   Kida H., Kawaoka Y., Naeve C.W., Webster R.G.;
RT   "Antigenic and genetic conservation of H3 influenza virus in wild
RT   ducks.";
RL   Virology 159:109-119(1987).
CC   -!- FUNCTION: HEMAGGLUTININ IS RESPONSIBLE FOR ATTACHING THE VIRUS TO
CC       CELL RECEPTORS AND FOR INITIATING INFECTION.
CC   -!- SUBUNIT: HOMOTRIMER. EACH OF THE MONOMERS IS FORMED BY TWO CHAINS
CC       (HA1 AND HA2) LINKED BY A DISULFIDE BOND.
CC   -!- SIMILARITY: BELONGS TO THE INFLUENZA HEMAGGLUTININ FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M16737; AAA43143.1; -.
DR   PIR; A27813; HMIV77.
DR   InterPro; IPR001364; Hemagglutn.
DR   Pfam; PF00509; Hemagglutinin; 1.
DR   ProDom; PD000225; Hemagglutn; 1.
KW   Envelope protein; Hemagglutinin; Glycoprotein.
FT   NON_TER       1      1
FT   CHAIN         1    328       HEMAGGLUTININ HA1 CHAIN.
FT   CHAIN       330    550       HEMAGGLUTININ HA2 CHAIN.
FT   CARBOHYD      8      8       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     22     22       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     38     38       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    165    165       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    285    285       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    483    483       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   550 AA;  61705 MW;  7E7ACFE716FC969A CRC64;


  Query Match             100.0%;  Score 55;  DB 1;  Length 550;
  Best Local Similarity   100.0%;  Pred. No. 0.023;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 YPYDVPDYA 9
           |||||||||
Db      98 YPYDVPDYA 106


RESULT   6
HEMA_IADH2
ID   HEMA_IADH2     STANDARD;      PRT;   550 AA.
AC   P12583; Q84011;
DT   01-OCT-1989 (Rel. 12, Created)
DT   01-APR-1990 (Rel. 14, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   HEMAGGLUTININ PRECURSOR [CONTAINS: HEMAGGLUTININ HA1 CHAIN;
DE   HEMAGGLUTININ HA2 CHAIN] (FRAGMENT).
GN   HA.
OS   Influenza A virus (strain A/Duck/Hokkaido/8/80).
OC   Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;
OC   Influenza virus A and B group; Influenza A viruses;
OC   Influenza A virus.
OX   NCBI_TaxID=11358;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=87265458; PubMed=2440178;
RA   Kida H., Kawaoka Y., Naeve C.W., Webster R.G.;
RT   "Antigenic and genetic conservation of H3 influenza virus in wild
RT   ducks.";
RL   Virology 159:109-119(1987).
CC   -!- FUNCTION: HEMAGGLUTININ IS RESPONSIBLE FOR ATTACHING THE VIRUS TO
CC       CELL RECEPTORS AND FOR INITIATING INFECTION.
CC   -!- SUBUNIT: HOMOTRIMER. EACH OF THE MONOMERS IS FORMED BY TWO CHAINS
CC       (HA1 AND HA2) LINKED BY A DISULFIDE BOND.
CC   -!- SIMILARITY: BELONGS TO THE INFLUENZA HEMAGGLUTININ FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M16738; AAA43144.1; -.
DR   PIR; B27813; HMIV80.
DR   InterPro; IPR001364; Hemagglutn.
DR   Pfam; PF00509; Hemagglutinin; 1.
DR   ProDom; PD000225; Hemagglutn; 1.
```

```
KW   Envelope protein; Hemagglutinin; Glycoprotein.
FT   NON_TER       1      1
FT   CHAIN         1    328       HEMAGGLUTININ HA1 CHAIN.
FT   CHAIN       330    550       HEMAGGLUTININ HA2 CHAIN.
FT   CARBOHYD      8      8       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     22     22       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     38     38       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    165    165       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    285    285       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    483    483       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CONFLICT    137    137       K -> N (IN PIR DATA BANK).
SQ   SEQUENCE   550 AA;  61659 MW;  A107023ACC9CC353 CRC64;


  Query Match             100.0%;  Score 55;  DB 1;  Length 550;
  Best Local Similarity   100.0%;  Pred. No. 0.023;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 YPYDVPDYA 9
           |||||||||
Db      98 YPYDVPDYA 106


RESULT   7
HEMA_IADH3
ID   HEMA_IADH3     STANDARD;      PRT;   550 AA.
AC   P12584; Q84012; Q89793;
DT   01-OCT-1989 (Rel. 12, Created)
DT   01-OCT-1989 (Rel. 12, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   HEMAGGLUTININ PRECURSOR [CONTAINS: HEMAGGLUTININ HA1 CHAIN;
DE   HEMAGGLUTININ HA2 CHAIN] (FRAGMENT).
GN   HA.
OS   Influenza A virus (strain A/Duck/Hokkaido/33/80).
OC   Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;
OC   Influenza virus A and B group; Influenza A viruses;
OC   Influenza A virus.
OX   NCBI_TaxID=11359;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=87265458; PubMed=2440178;
RA   Kida H., Kawaoka Y., Naeve C.W., Webster R.G.;
RT   "Antigenic and genetic conservation of H3 influenza virus in wild
RT   ducks.";
RL   Virology 159:109-119(1987).
CC   -!- FUNCTION: HEMAGGLUTININ IS RESPONSIBLE FOR ATTACHING THE VIRUS TO
CC       CELL RECEPTORS AND FOR INITIATING INFECTION.
CC   -!- SUBUNIT: HOMOTRIMER. EACH OF THE MONOMERS IS FORMED BY TWO CHAINS
CC       (HA1 AND HA2) LINKED BY A DISULFIDE BOND.
CC   -!- SIMILARITY: BELONGS TO THE INFLUENZA HEMAGGLUTININ FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M16739; AAA43145.1; -.
DR   PIR; C27813; HMIV33.
DR   InterPro; IPR001364; Hemagglutn.
DR   Pfam; PF00509; Hemagglutinin; 1.
DR   ProDom; PD000225; Hemagglutn; 1.
KW   Envelope protein; Hemagglutinin; Glycoprotein.
FT   NON_TER       1      1
FT   CHAIN         1    328       HEMAGGLUTININ HA1 CHAIN.
FT   CHAIN       330    550       HEMAGGLUTININ HA2 CHAIN.
FT   CARBOHYD      8      8       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     22     22       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     38     38       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    165    165       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    285    285       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    483    483       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   550 AA;  61577 MW;  6C30BF67CFDCB7DE CRC64;


  Query Match             100.0%;  Score 55;  DB 1;  Length 550;
  Best Local Similarity   100.0%;  Pred. No. 0.023;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 YPYDVPDYA 9
           |||||||||
Db      98 YPYDVPDYA 106


RESULT   8
HEMA_IADH4
ID   HEMA_IADH4     STANDARD;      PRT;   550 AA.
AC   P12585; Q84013; Q84014;
DT   01-OCT-1989 (Rel. 12, Created)
DT   01-OCT-1989 (Rel. 12, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   HEMAGGLUTININ PRECURSOR [CONTAINS: HEMAGGLUTININ HA1 CHAIN;
DE   HEMAGGLUTININ HA2 CHAIN] (FRAGMENT).
GN   HA.
OS   Influenza A virus (strain A/Duck/Hokkaido/7/82).
OC   Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;
OC   Influenza virus A and B group; Influenza A viruses;
OC   Influenza A virus.
OX   NCBI_TaxID=11360;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=87265458; PubMed=2440178;
RA   Kida H., Kawaoka Y., Naeve C.W., Webster R.G.;
RT   "Antigenic and genetic conservation of H3 influenza virus in wild
RT   ducks.";
RL   Virology 159:109-119(1987).
CC   -!- FUNCTION: HEMAGGLUTININ IS RESPONSIBLE FOR ATTACHING THE VIRUS TO
CC       CELL RECEPTORS AND FOR INITIATING INFECTION.
CC   -!- SUBUNIT: HOMOTRIMER. EACH OF THE MONOMERS IS FORMED BY TWO CHAINS
CC       (HA1 AND HA2) LINKED BY A DISULFIDE BOND.
CC   -!- SIMILARITY: BELONGS TO THE INFLUENZA HEMAGGLUTININ FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M16740; AAA43146.1; -.
DR   PIR; D27813; HMIV89.
DR   InterPro; IPR001364; Hemagglutn.
DR   Pfam; PF00509; Hemagglutinin; 1.
DR   ProDom; PD000225; Hemagglutn; 1.
KW   Envelope protein; Hemagglutinin; Glycoprotein.
FT   NON_TER       1      1
FT   CHAIN         1    328       HEMAGGLUTININ HA1 CHAIN.
FT   CHAIN       330    550       HEMAGGLUTININ HA2 CHAIN.
FT   CARBOHYD      8      8       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     22     22       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     38     38       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    165    165       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    285    285       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    483    483       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   550 AA;  61664 MW;  A16B2CF8CBBBD9D0 CRC64;


  Query Match             100.0%;  Score 55;  DB 1;  Length 550;
  Best Local Similarity   100.0%;  Pred. No. 0.023;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 YPYDVPDYA 9
           |||||||||
```

```
Db     98 YPYDVPDYA 106


RESULT   9
HEMA_IADH6
ID   HEMA_IADH6     STANDARD;      PRT;   550 AA.
AC   P12587; Q84017;
DT   01-OCT-1989 (Rel. 12, Created)
DT   01-APR-1990 (Rel. 14, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   HEMAGGLUTININ PRECURSOR [CONTAINS: HEMAGGLUTININ HA1 CHAIN;
DE   HEMAGGLUTININ HA2 CHAIN] (FRAGMENT).
GN   HA.
OS   Influenza A virus (strain A/Duck/Hokkaido/9/85).
OC   Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;
OC   Influenza virus A and B group; Influenza A viruses;
OC   Influenza A virus.
OX   NCBI_TaxID=11362;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=87265458; PubMed=2440178;
RA   Kida H., Kawaoka Y., Naeve C.W., Webster R.G.;
RT   "Antigenic and genetic conservation of H3 influenza virus in wild
RT   ducks.";
RL   Virology 159:109-119(1987).
CC   -!- FUNCTION: HEMAGGLUTININ IS RESPONSIBLE FOR ATTACHING THE VIRUS TO
CC       CELL RECEPTORS AND FOR INITIATING INFECTION.
CC   -!- SUBUNIT: HOMOTRIMER. EACH OF THE MONOMERS IS FORMED BY TWO CHAINS
CC       (HA1 AND HA2) LINKED BY A DISULFIDE BOND.
CC   -!- SIMILARITY: BELONGS TO THE INFLUENZA HEMAGGLUTININ FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M16742; AAA43148.1; -.
DR   PIR; F27813; HMIV98.
DR   InterPro; IPR001364; Hemagglutn.
DR   Pfam; PF00509; Hemagglutinin; 1.
DR   ProDom; PD000225; Hemagglutn; 1.
KW   Envelope protein; Hemagglutinin; Glycoprotein.
FT   NON_TER       1      1
FT   CHAIN         1    328       HEMAGGLUTININ HA1 CHAIN.
FT   CHAIN       330    550       HEMAGGLUTININ HA2 CHAIN.
FT   CARBOHYD     22     22       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     38     38       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    165    165       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    285    285       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    483    483       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CONFLICT      8      8       Y -> N (IN PIR DATA BANK).
SQ   SEQUENCE   550 AA;  61711 MW;  67BCD85F44736CFE CRC64;


  Query Match             100.0%;  Score 55;  DB 1;  Length 550;
  Best Local Similarity   100.0%;  Pred. No. 0.023;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 YPYDVPDYA 9
           |||||||||
Db      98 YPYDVPDYA 106


RESULT  10
HEMA_IADH7
ID   HEMA_IADH7     STANDARD;      PRT;   550 AA.
AC   P12588; Q84018; Q89470;
DT   01-OCT-1989 (Rel. 12, Created)
DT   01-OCT-1989 (Rel. 12, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   HEMAGGLUTININ PRECURSOR [CONTAINS: HEMAGGLUTININ HA1 CHAIN;
DE   HEMAGGLUTININ HA2 CHAIN] (FRAGMENT).
GN   HA.
OS   Influenza A virus (strain A/Duck/Hokkaido/10/85).
OC   Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;
OC   Influenza virus A and B group; Influenza A viruses;
OC   Influenza A virus.
OX   NCBI_TaxID=11363;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=87265458; PubMed=2440178;
RA   Kida H., Kawaoka Y., Naeve C.W., Webster R.G.;
RT   "Antigenic and genetic conservation of H3 influenza virus in wild
RT   ducks.";
RL   Virology 159:109-119(1987).
CC   -!- FUNCTION: HEMAGGLUTININ IS RESPONSIBLE FOR ATTACHING THE VIRUS TO
CC       CELL RECEPTORS AND FOR INITIATING INFECTION.
CC   -!- SUBUNIT: HOMOTRIMER. EACH OF THE MONOMERS IS FORMED BY TWO CHAINS
CC       (HA1 AND HA2) LINKED BY A DISULFIDE BOND.
CC   -!- SIMILARITY: BELONGS TO THE INFLUENZA HEMAGGLUTININ FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M16743; AAA43149.1; -.
DR   PIR; G27813; HMIV15.
DR   InterPro; IPR001364; Hemagglutn.
DR   Pfam; PF00509; Hemagglutinin; 1.
DR   ProDom; PD000225; Hemagglutn; 1.
KW   Envelope protein; Hemagglutinin; Glycoprotein.
FT   NON_TER       1      1
FT   CHAIN         1    328       HEMAGGLUTININ HA1 CHAIN.
FT   CHAIN       330    550       HEMAGGLUTININ HA2 CHAIN.
FT   CARBOHYD      8      8       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     22     22       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     38     38       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    165    165       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    285    285       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    483    483       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   550 AA;  61761 MW;  6EF81793281D53EB CRC64;


  Query Match             100.0%;  Score 55;  DB 1;  Length 550;
  Best Local Similarity   100.0%;  Pred. No. 0.023;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 YPYDVPDYA 9
           |||||||||
Db      98 YPYDVPDYA 106


RESULT  11
HEMA_IADHK
ID   HEMA_IADHK     STANDARD;      PRT;   550 AA.
AC   P43257;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   HEMAGGLUTININ PRECURSOR [CONTAINS: HEMAGGLUTININ HA1 CHAIN;
DE   HEMAGGLUTININ HA2 CHAIN] (FRAGMENT).
GN   HA.
OS   Influenza A virus (strain A/Duck/Hong Kong/7/75).
OC   Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;
OC   Influenza virus A and B group; Influenza A viruses;
OC   Influenza A virus.
OX   NCBI_TaxID=11364;
RN   [1]
RP   SEQUENCE FROM N.A.
```

```
RX   MEDLINE=91341491; PubMed=1875195;
RA   Yasuda J., Shortridge K.F., Shimizu Y., Kida H.;
RT   "Molecular evidence for a role of domestic ducks in the introduction
RT   of avian H3 influenza viruses to pigs in southern China, where the
RT   A/Hong Kong/68 (H3N2) strain emerged.";
RL   J. Gen. Virol. 72:2007-2010(1991).
CC   -!- FUNCTION: HEMAGGLUTININ IS RESPONSIBLE FOR ATTACHING THE VIRUS TO
CC       CELL RECEPTORS AND FOR INITIATING INFECTION.
CC   -!- SUBUNIT: HOMOTRIMER. EACH OF THE MONOMERS IS FORMED BY TWO CHAINS
CC       (HA1 AND HA2) LINKED BY A DISULFIDE BOND.
CC   -!- SIMILARITY: BELONGS TO THE INFLUENZA HEMAGGLUTININ FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; D00929; BAA00769.1; -.
DR   HSSP; P03437; 5HMG.
DR   InterPro; IPR001364; Hemagglutn.
DR   Pfam; PF00509; Hemagglutinin; 1.
DR   ProDom; PD000225; Hemagglutn; 1.
KW   Envelope protein; Hemagglutinin; Glycoprotein.
FT   NON_TER       1      1
FT   CHAIN         1    328       HEMAGGLUTININ HA1 CHAIN.
FT   CHAIN       330    550       HEMAGGLUTININ HA2 CHAIN.
FT   CARBOHYD      8      8       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     22     22       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     38     38       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    165    165       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    285    285       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    483    483       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   550 AA;  61549 MW;  864639B829FE1BA9 CRC64;


  Query Match             100.0%;  Score 55;  DB 1;  Length 550;
  Best Local Similarity   100.0%;  Pred. No. 0.023;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 YPYDVPDYA 9
           |||||||||
Db      98 YPYDVPDYA 106


RESULT  12
HEMA_IADHL
ID   HEMA_IADHL     STANDARD;      PRT;   550 AA.
AC   P43258;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   HEMAGGLUTININ PRECURSOR [CONTAINS: HEMAGGLUTININ HA1 CHAIN;
DE   HEMAGGLUTININ HA2 CHAIN] (FRAGMENT).
GN   HA.
OS   Influenza A virus (strain A/Duck/Hong Kong/64/76).
OC   Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;
OC   Influenza virus A and B group; Influenza A viruses;
OC   Influenza A virus.
OX   NCBI_TaxID=45412;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=91341491; PubMed=1875195;
RA   Yasuda J., Shortridge K.F., Shimizu Y., Kida H.;
RT   "Molecular evidence for a role of domestic ducks in the introduction
RT   of avian H3 influenza viruses to pigs in southern China, where the
RT   A/Hong Kong/68 (H3N2) strain emerged.";
RL   J. Gen. Virol. 72:2007-2010(1991).
CC   -!- FUNCTION: HEMAGGLUTININ IS RESPONSIBLE FOR ATTACHING THE VIRUS TO
CC       CELL RECEPTORS AND FOR INITIATING INFECTION.
CC   -!- SUBUNIT: HOMOTRIMER. EACH OF THE MONOMERS IS FORMED BY TWO CHAINS
CC       (HA1 AND HA2) LINKED BY A DISULFIDE BOND.
CC   -!- SIMILARITY: BELONGS TO THE INFLUENZA HEMAGGLUTININ FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; D00931; BAA00771.1; -.
DR   HSSP; P03437; 2HMG.
DR   InterPro; IPR001364; Hemagglutn.
DR   Pfam; PF00509; Hemagglutinin; 1.
DR   ProDom; PD000225; Hemagglutn; 1.
KW   Envelope protein; Hemagglutinin; Glycoprotein.
FT   NON_TER       1      1
FT   CHAIN         1    328       HEMAGGLUTININ HA1 CHAIN.
FT   CHAIN       330    550       HEMAGGLUTININ HA2 CHAIN.
FT   CARBOHYD      8      8       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     22     22       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     38     38       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    165    165       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    285    285       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    483    483       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   550 AA;  61718 MW;  A351C56789E4BE9A CRC64;


  Query Match             100.0%;  Score 55;  DB 1;  Length 550;
  Best Local Similarity   100.0%;  Pred. No. 0.023;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 YPYDVPDYA 9
           |||||||||
Db      98 YPYDVPDYA 106


RESULT  13
HEMA_IADHM
ID   HEMA_IADHM     STANDARD;      PRT;   550 AA.
AC   P43259;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   HEMAGGLUTININ PRECURSOR [CONTAINS: HEMAGGLUTININ HA1 CHAIN;
DE   HEMAGGLUTININ HA2 CHAIN] (FRAGMENT).
GN   HA.
OS   Influenza A virus (strain A/Duck/Hong Kong/231/77).
OC   Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;
OC   Influenza virus A and B group; Influenza A viruses;
OC   Influenza A virus.
OX   NCBI_TaxID=45411;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=91341491; PubMed=1875195;
RA   Yasuda J., Shortridge K.F., Shimizu Y., Kida H.;
RT   "Molecular evidence for a role of domestic ducks in the introduction
RT   of avian H3 influenza viruses to pigs in southern China, where the
RT   A/Hong Kong/68 (H3N2) strain emerged.";
RL   J. Gen. Virol. 72:2007-2010(1991).
CC   -!- FUNCTION: HEMAGGLUTININ IS RESPONSIBLE FOR ATTACHING THE VIRUS TO
CC       CELL RECEPTORS AND FOR INITIATING INFECTION.
CC   -!- SUBUNIT: HOMOTRIMER. EACH OF THE MONOMERS IS FORMED BY TWO CHAINS
CC       (HA1 AND HA2) LINKED BY A DISULFIDE BOND.
CC   -!- SIMILARITY: BELONGS TO THE INFLUENZA HEMAGGLUTININ FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
```

```
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; D00932; BAA00772.1; -.
DR   InterPro; IPR001364; Hemagglutn.
DR   Pfam; PF00509; Hemagglutinin; 1.
DR   ProDom; PD000225; Hemagglutn; 1.
KW   Envelope protein; Hemagglutinin; Glycoprotein.
FT   NON_TER       1      1
FT   CHAIN         1    328       HEMAGGLUTININ HA1 CHAIN.
FT   CHAIN       330    550       HEMAGGLUTININ HA2 CHAIN.
FT   CARBOHYD     22     22       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     38     38       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    165    165       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    285    285       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    483    483       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   550 AA;  61762 MW;  6FEF8B49488C191A CRC64;


  Query Match             100.0%;  Score 55;  DB 1;  Length 550;
  Best Local Similarity   100.0%;  Pred. No. 0.023;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 YPYDVPDYA 9
          |||||||||
Db     98 YPYDVPDYA 106


RESULT  14
HEMA_IAGHK
ID   HEMA_IAGHK     STANDARD;      PRT;   550 AA.
AC   P43260;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   HEMAGGLUTININ PRECURSOR [CONTAINS: HEMAGGLUTININ HA1 CHAIN;
DE   HEMAGGLUTININ HA2 CHAIN] (FRAGMENT).
GN   HA.
OS   Influenza A virus (strain A/Goose/Hong Kong/10/76).
OC   Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;
OC   Influenza virus A and B group; Influenza A viruses;
OC   Influenza A virus.
OX   NCBI_TaxID=45414;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=91341491; PubMed=1875195;
RA   Yasuda J., Shortridge K.F., Shimizu Y., Kida H.;
RT   "Molecular evidence for a role of domestic ducks in the introduction
RT   of avian H3 influenza viruses to pigs in southern China, where the
RT   A/Hong Kong/68 (H3N2) strain emerged.";
RL   J. Gen. Virol. 72:2007-2010(1991).
CC   -!- FUNCTION: HEMAGGLUTININ IS RESPONSIBLE FOR ATTACHING THE VIRUS TO
CC       CELL RECEPTORS AND FOR INITIATING INFECTION.
CC   -!- SUBUNIT: HOMOTRIMER. EACH OF THE MONOMERS IS FORMED BY TWO CHAINS
CC       (HA1 AND HA2) LINKED BY A DISULFIDE BOND.
CC   -!- SIMILARITY: BELONGS TO THE INFLUENZA HEMAGGLUTININ FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; D00930; BAA00770.1; -.
DR   HSSP; P03437; 5HMG.
DR   InterPro; IPR001364; Hemagglutn.
DR   Pfam; PF00509; Hemagglutinin; 1.
DR   ProDom; PD000225; Hemagglutn; 1.
KW   Envelope protein; Hemagglutinin; Glycoprotein.
FT   NON_TER       1      1
FT   CHAIN         1    328       HEMAGGLUTININ HA1 CHAIN.
FT   CHAIN       330    550       HEMAGGLUTININ HA2 CHAIN.
FT   CARBOHYD      8      8       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     22     22       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     38     38       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    165    165       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    285    285       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    483    483       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   550 AA;  61676 MW;  9A1E094DA28BACD2 CRC64;


  Query Match             100.0%;  Score 55;  DB 1;  Length 550;
  Best Local Similarity   100.0%;  Pred. No. 0.023;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 YPYDVPDYA 9
          |||||||||
Db     98 YPYDVPDYA 106


RESULT  15
HEMA_IAME6
ID   HEMA_IAME6     STANDARD;      PRT;   550 AA.
AC   P12589;
DT   01-OCT-1989 (Rel. 12, Created)
DT   01-OCT-1989 (Rel. 12, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   HEMAGGLUTININ PRECURSOR [CONTAINS: HEMAGGLUTININ HA1 CHAIN;
DE   HEMAGGLUTININ HA2 CHAIN] (FRAGMENT).
GN   HA.
OS   Influenza A virus (strain A/Memphis/6/86).
OC   Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;
OC   Influenza virus A and B group; Influenza A viruses;
OC   Influenza A virus.
OX   NCBI_TaxID=11440;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=88306236; PubMed=3407150;
RA   Katz J.M., Webster R.G.;
RT   "Antigenic and structural characterization of multiple subpopulations
RT   of H3N2 influenza virus from an individual.";
RL   Virology 165:446-456(1988).
CC   -!- FUNCTION: HEMAGGLUTININ IS RESPONSIBLE FOR ATTACHING THE VIRUS TO
CC       CELL RECEPTORS AND FOR INITIATING INFECTION.
CC   -!- SUBUNIT: HOMOTRIMER. EACH OF THE MONOMERS IS FORMED BY TWO CHAINS
CC       (HA1 AND HA2) LINKED BY A DISULFIDE BOND.
CC   -!- SIMILARITY: BELONGS TO THE INFLUENZA HEMAGGLUTININ FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M21648; AAA43275.1; -.
DR   PIR; A29245; HMIV86.
DR   HSSP; P03437; 2HMG.
DR   InterPro; IPR001364; Hemagglutn.
DR   Pfam; PF00509; Hemagglutinin; 1.
DR   ProDom; PD000225; Hemagglutn; 1.
KW   Envelope protein; Hemagglutinin; Glycoprotein.
FT   NON_TER       1      1
FT   CHAIN         1    328       HEMAGGLUTININ HA1 CHAIN.
FT   CHAIN       330    550       HEMAGGLUTININ HA2 CHAIN.
FT   CARBOHYD      8      8       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     22     22       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     38     38       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     63     63       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    126    126       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    165    165       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    246    246       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    285    285       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    483    483       N-LINKED (GLCNAC. . .) (POTENTIAL).
```

SQ SEQUENCE 550 AA; 61804 MW; 52C9F14B309310ED CRC64;

Query Match 100.0%; Score 55; DB 1; Length 550;
Best Local Similarity 100.0%; Pred. No. 0.023;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       1 YPYDVPDYA 9
           |||||||||
Db      98 YPYDVPDYA 106
```

Search completed: April 11, 2002, 10:22:00
Job time: 430 sec

GenCore version 4.5

OM protein - protein search, using sw model

Run on: April 11, 2002, 10:15:50 ; Search time 50.04 Seconds
(without alignments)
26.308 Million cell updates/sec

Title: US-09-284-787-1
Perfect score: 55
Sequence: 1 YPYDVPDYA 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 473505 seqs, 146272329 residues

Total number of hits satisfying chosen parameters: 473505

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_17:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	55	100.0	236	12	Q9W8Z1	Q9w8z1 influenza a
2	55	100.0	236	12	Q9W8T9	Q9w8t9 influenza a
3	55	100.0	236	12	Q9W8J7	Q9w8j7 influenza a
4	55	100.0	286	12	Q9WI26	Q9wi26 influenza a
5	55	100.0	286	12	Q9WI25	Q9wi25 influenza a
6	55	100.0	286	12	Q9WI24	Q9wi24 influenza a
7	55	100.0	325	12	O40774	O40774 influenza a
8	55	100.0	325	12	Q9DL45	Q9dl45 influenza a
9	55	100.0	326	12	O40779	O40779 influenza a
10	55	100.0	326	12	O40780	O40780 influenza a
11	55	100.0	327	12	O92324	O92324 influenza a
12	55	100.0	327	12	O92325	O92325 influenza a
13	55	100.0	327	12	O92327	O92327 influenza a
14	55	100.0	327	12	Q9YT90	Q9yt90 influenza a
15	55	100.0	327	12	Q9YT89	Q9yt89 influenza a
16	55	100.0	327	12	Q9YT88	Q9yt88 influenza a
17	55	100.0	327	12	Q9YT87	Q9yt87 influenza a
18	55	100.0	327	12	Q9YT86	Q9yt86 influenza a
19	55	100.0	327	12	Q9YRU9	Q9yru9 influenza a
20	55	100.0	327	12	Q9YRU8	Q9yru8 influenza a
21	55	100.0	327	12	Q9YRU7	Q9yru7 influenza a
22	55	100.0	327	12	Q9YRU6	Q9yru6 influenza a
23	55	100.0	327	12	Q9YRU5	Q9yru5 influenza a
24	55	100.0	327	12	Q9YRU4	Q9yru4 influenza a
25	55	100.0	327	12	Q9YRU3	Q9yru3 influenza a
26	55	100.0	327	12	Q9YJ57	Q9yj57 influenza a
27	55	100.0	327	12	Q9YIN9	Q9yin9 influenza a
28	55	100.0	327	12	Q9Q884	Q9q884 influenza a
29	55	100.0	328	12	O40855	O40855 influenza a
30	55	100.0	328	12	O40867	O40867 influenza a
31	55	100.0	328	12	Q82567	Q82567 influenza a
32	55	100.0	328	12	Q82568	Q82568 influenza a
33	55	100.0	328	12	Q82569	Q82569 influenza a
34	55	100.0	328	12	Q82578	Q82578 influenza a
35	55	100.0	328	12	Q82579	Q82579 influenza a
36	55	100.0	328	12	Q82580	Q82580 influenza a
37	55	100.0	328	12	Q82581	Q82581 influenza a
38	55	100.0	328	12	Q82575	Q82575 influenza a
39	55	100.0	328	12	Q82576	Q82576 influenza a
40	55	100.0	328	12	Q82577	Q82577 influenza a
41	55	100.0	328	12	Q82582	Q82582 influenza a
42	55	100.0	328	12	Q82583	Q82583 influenza a
43	55	100.0	328	12	Q82586	Q82586 influenza a
44	55	100.0	328	12	Q82587	Q82587 influenza a
45	55	100.0	328	12	Q82589	Q82589 influenza a

ALIGNMENTS

```
RESULT   1
Q9W8Z1
ID   Q9W8Z1      PRELIMINARY;      PRT;   236 AA.
AC   Q9W8Z1;
DT   01-NOV-1999 (TrEMBLrel. 12, Created)
DT   01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   HEMAGGLUTININ (FRAGMENT).
OS   Influenza A virus.
OC   Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;
OC   Influenza virus A and B group; Influenza A viruses.
OX   NCBI_TaxID=11320;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=TOKYO1527, TOKYO1511;
RA   Mori S., Sekine H.;
RT   "HA1 domain of Influenza A (H3N2) virus.";
RL   Submitted (MAY-1998) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AB013807; BAA77285.1; -.
DR   EMBL; AB013806; BAA77284.1; -.
DR   InterPro; IPR001364; Hemagglutn.
DR   Pfam; PF00509; Hemagglutinin; 1.
DR   PRINTS; PR00329; HEMAGGLUTN12.
DR   ProDom; PD000225; Hemagglutn; 1.
FT   NON_TER       1      1
FT   NON_TER     236    236
SQ   SEQUENCE   236 AA;  26477 MW;  4566C8E7210FB558 CRC64;


  Query Match            100.0%;  Score 55;  DB 12;  Length 236;
  Best Local Similarity  100.0%;  Pred. No. 0.05;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 YPYDVPDYA 9
           |||||||||
Db      11 YPYDVPDYA 19


RESULT   2
Q9W8T9
ID   Q9W8T9      PRELIMINARY;      PRT;   236 AA.
AC   Q9W8T9;
```

```
DT   01-NOV-1999 (TrEMBLrel. 12, Created)
DT   01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   HEMAGGLUTININ (FRAGMENT).
OS   Influenza A virus.
OC   Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;
OC   Influenza virus A and B group; Influenza A viruses.
OX   NCBI_TaxID=11320;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=TOKYO1570, TOKYO1566, TOKYO1567, TOKYO1568, TOKYO1569;
RA   Mori S., Sekine H.;
RT   "HA1 domain of Influenza A (H3N2) virus.";
RL   Submitted (MAY-1998) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AB013813; BAA77291.1; -.
DR   EMBL; AB013809; BAA77287.1; -.
DR   EMBL; AB013810; BAA77288.1; -.
DR   EMBL; AB013811; BAA77289.1; -.
DR   EMBL; AB013812; BAA77290.1; -.
DR   InterPro; IPR001364; Hemagglutn.
DR   Pfam; PF00509; Hemagglutinin; 1.
DR   PRINTS; PR00329; HEMAGGLUTN12.
DR   ProDom; PD000225; Hemagglutn; 1.
FT   NON_TER       1      1
FT   NON_TER     236    236
SQ   SEQUENCE   236 AA;  26491 MW;  44CDD257210FB558 CRC64;


  Query Match             100.0%;  Score 55;  DB 12;  Length 236;
  Best Local Similarity   100.0%;  Pred. No. 0.05;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 YPYDVPDYA 9
          |||||||||
Db     11 YPYDVPDYA 19


RESULT   3
Q9W8J7
ID   Q9W8J7         PRELIMINARY;   PRT;   236 AA.
AC   Q9W8J7;
DT   01-NOV-1999 (TrEMBLrel. 12, Created)
DT   01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   HEMAGGLUTININ (FRAGMENT).
GN   HA.
OS   Influenza A virus.
OC   Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;
OC   Influenza virus A and B group; Influenza A viruses.
OX   NCBI_TaxID=11320;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=72, 70, AND 71;
RA   Nagashima M., Mori S., Sekine H.;
RT   "Influenza A virus gene for hemagglutinin.";
RL   Submitted (MAY-1998) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=TOKYO1539;
RA   Mori S., Sekine H.;
RT   "HA1 domain of Influenza A (H3N2) virus.";
RL   Submitted (MAY-1998) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AB014062; BAA77294.1; -.
DR   EMBL; AB013808; BAA77286.1; -.
DR   EMBL; AB014060; BAA77292.1; -.
DR   EMBL; AB014061; BAA77293.1; -.
DR   InterPro; IPR001364; Hemagglutn.
DR   Pfam; PF00509; Hemagglutinin; 1.
DR   PRINTS; PR00329; HEMAGGLUTN12.
DR   ProDom; PD000225; Hemagglutn; 1.
FT   NON_TER       1      1
FT   NON_TER     236    236
SQ   SEQUENCE   236 AA;  26493 MW;  457455682C8D28D7 CRC64;


  Query Match             100.0%;  Score 55;  DB 12;  Length 236;
  Best Local Similarity   100.0%;  Pred. No. 0.05;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 YPYDVPDYA 9
          |||||||||
Db     11 YPYDVPDYA 19


RESULT   4
Q9WI26
ID   Q9WI26      PRELIMINARY;      PRT;   286 AA.
AC   Q9WI26;
DT   01-NOV-1999 (TrEMBLrel. 12, Created)
DT   01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   H3 HEMAGGLUTININ (FRAGMENT).
GN   HA.
OS   influenza A virus (A/Taiwan/2034/96(H3N2)).
OC   Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;
OC   Influenza virus A and B group; Influenza A viruses; Influenza A virus.
OX   NCBI_TaxID=95241;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=A/TAIWAN/2034/96;
RA   Shih S.-R., Chen S.-H., Wu S.-L., Ho C.-C.;
RT   "Sequence Analysis of HA and NA Genes of Influenza Viruses Isolated in
RT   Taiwan.";
RL   Submitted (APR-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF139937; AAD34854.1; -.
DR   InterPro; IPR001364; Hemagglutn.
DR   Pfam; PF00509; Hemagglutinin; 1.
DR   PRINTS; PR00329; HEMAGGLUTN12.
DR   ProDom; PD000225; Hemagglutn; 1.
FT   NON_TER       1      1
FT   NON_TER     286    286
SQ   SEQUENCE   286 AA;  32075 MW;  7BA39C7632D33186 CRC64;


  Query Match             100.0%;  Score 55;  DB 12;  Length 286;
  Best Local Similarity   100.0%;  Pred. No. 0.063;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 YPYDVPDYA 9
          |||||||||
Db     57 YPYDVPDYA 65


RESULT   5
Q9WI25
ID   Q9WI25      PRELIMINARY;      PRT;   286 AA.
AC   Q9WI25;
DT   01-NOV-1999 (TrEMBLrel. 12, Created)
DT   01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   H3 HEMAGGLUTININ (FRAGMENT).
GN   HA.
OS   influenza A virus (A/Taiwan/2192/96(H3N2)).
OC   Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;
OC   Influenza virus A and B group; Influenza A viruses; Influenza A virus.
OX   NCBI_TaxID=95240;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=A/TAIWAN/2192/96;
RA   Shih S.-R., Chen S.-H., Wu S.-L., Ho C.-C.;
RT   "Sequence Analysis of HA and NA Genes of Influenza Viruses Isolated in
RT   Taiwan.";
RL   Submitted (APR-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF139936; AAD34853.1; -.
DR   InterPro; IPR001364; Hemagglutn.
DR   Pfam; PF00509; Hemagglutinin; 1.
```

```
DR   PRINTS; PR00329; HEMAGGLUTN12.
DR   ProDom; PD000225; Hemagglutn; 1.
FT   NON_TER       1      1
FT   NON_TER     286    286
SQ   SEQUENCE   286 AA;  32059 MW;  CB51B092AAD0D7D8 CRC64;


  Query Match             100.0%;  Score 55;  DB 12;  Length 286;
  Best Local Similarity   100.0%;  Pred. No. 0.063;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 YPYDVPDYA 9
          |||||||||
Db     57 YPYDVPDYA 65


RESULT   6
Q9WI24
ID   Q9WI24      PRELIMINARY;      PRT;   286 AA.
AC   Q9WI24;
DT   01-NOV-1999 (TrEMBLrel. 12, Created)
DT   01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   H3 HEMAGGLUTININ (FRAGMENT).
GN   HA.
OS   Influenza A virus (A/Taiwan/2191/96(H3N2)).
OC   Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;
OC   Influenza virus A and B group; Influenza A viruses; Influenza A virus.
OX   NCBI_TaxID=95236;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=A/TAIWAN/2191/96;
RA   Shih S.-R., Chen S.-H., Wu S.-L., Ho C.-C.;
RT   "Sequence Analysis of HA and NA Genes of Influenza Viruses Isolated in
RT   Taiwan.";
RL   Submitted (APR-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF139932; AAD34849.1; -.
DR   InterPro; IPR001364; Hemagglutn.
DR   Pfam; PF00509; Hemagglutinin; 1.
DR   PRINTS; PR00329; HEMAGGLUTN12.
DR   ProDom; PD000225; Hemagglutn; 1.
FT   NON_TER       1      1
FT   NON_TER     286    286
SQ   SEQUENCE   286 AA;  31976 MW;  69AB2AE727D26078 CRC64;


  Query Match             100.0%;  Score 55;  DB 12;  Length 286;
  Best Local Similarity   100.0%;  Pred. No. 0.063;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 YPYDVPDYA 9
          |||||||||
Db     57 YPYDVPDYA 65


RESULT   7
O40774
ID   O40774      PRELIMINARY;      PRT;   325 AA.
AC   O40774;
DT   01-JAN-1998 (TrEMBLrel. 05, Created)
DT   01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   HEMAGGLUTININ GENE (FRAGMENT).
GN   HA.
OS   Influenza A virus.
OC   Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;
OC   Influenza virus A and B group; Influenza A viruses.
OX   NCBI_TaxID=11320;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=A/ILLINOIS/5/95(H3N2);
RA   Fitch W.M., Bush R.M., Bender C.A., Cox N.J.;
RL   Submitted (JUN-1997) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF008782; AAB63719.1; -.
DR   InterPro; IPR001364; Hemagglutn.
DR   Pfam; PF00509; Hemagglutinin; 1.
DR   ProDom; PD000225; Hemagglutn; 1.
FT   NON_TER       1      1
FT   NON_TER     325    325
SQ   SEQUENCE   325 AA;  36103 MW;  4E5B4648C2024765 CRC64;


  Query Match             100.0%;  Score 55;  DB 12;  Length 325;
  Best Local Similarity   100.0%;  Pred. No. 0.073;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 YPYDVPDYA 9
          |||||||||
Db     98 YPYDVPDYA 106


RESULT   8
Q9DL45
ID   Q9DL45      PRELIMINARY;      PRT;   325 AA.
AC   Q9DL45;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   HEMAGGLUTININ (FRAGMENT).
OS   Influenza A virus (A/Finland/583/98(H3N2)).
OC   Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;
OC   Influenza virus A and B group; Influenza A viruses; Influenza A virus.
OX   NCBI_TaxID=147149;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=A/FINLAND/583/98;
RA   Pyhala R., Haanpaa M., Kleemola M., Tervahauta R., Visakorpi R.,
RA   Kinnunen L.;
RT   "Acceptable protective efficacy of influenza vaccination in young
RT   military conscripts in circumstances of incomplete antigenic and
RT   genetic match.";
RL   Vaccine 0:0-0(2001).
DR   EMBL; AF311689; AAG47810.1; -.
DR   InterPro; IPR001364; Hemagglutn.
DR   Pfam; PF00509; Hemagglutinin; 1.
DR   PRINTS; PR00329; HEMAGGLUTN12.
DR   ProDom; PD000225; Hemagglutn; 1.
FT   NON_TER       1      1
FT   NON_TER     325    325
SQ   SEQUENCE   325 AA;  36029 MW;  5BECB8CC3E83B1C2 CRC64;


  Query Match             100.0%;  Score 55;  DB 12;  Length 325;
  Best Local Similarity   100.0%;  Pred. No. 0.073;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 YPYDVPDYA 9
          |||||||||
Db     98 YPYDVPDYA 106


RESULT   9
O40779
ID   O40779      PRELIMINARY;      PRT;   326 AA.
AC   O40779;
DT   01-JAN-1998 (TrEMBLrel. 05, Created)
DT   01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   HEMAGGLUTININ GENE (FRAGMENT).
GN   HA.
OS   Influenza A virus.
OC   Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;
OC   Influenza virus A and B group; Influenza A viruses.
OX   NCBI_TaxID=11320;
RN   [1]
RP   SEQUENCE FROM N.A.
```

```
RC   STRAIN=A/NEW_YORK/28/94(H3N2);
RA   Fitch W.M., Bush R.M., Bender C.A., Cox N.J.;
RL   Submitted (JUN-1997) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF008787; AAB63724.1; -.
DR   InterPro; IPR001364; Hemagglutn.
DR   Pfam; PF00509; Hemagglutinin; 1.
DR   ProDom; PD000225; Hemagglutn; 1.
FT   NON_TER       1      1
FT   NON_TER     326    326
SQ   SEQUENCE   326 AA;  36197 MW;  4B84BD4AADC44A77 CRC64;


  Query Match             100.0%;  Score 55;  DB 12;  Length 326;
  Best Local Similarity   100.0%;  Pred. No. 0.074;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 YPYDVPDYA 9
          |||||||||
Db     98 YPYDVPDYA 106


RESULT  10
O40780
ID   O40780      PRELIMINARY;      PRT;   326 AA.
AC   O40780;
DT   01-JAN-1998 (TrEMBLrel. 05, Created)
DT   01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   HEMAGGLUTININ GENE (FRAGMENT).
GN   HA.
OS   Influenza A virus.
OC   Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;
OC   Influenza virus A and B group; Influenza A viruses.
OX   NCBI_TaxID=11320;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=A/PENNSYLVANIA/7/94(H3N2);
RA   Fitch W.M., Bush R.M., Bender C.A., Cox N.J.;
RL   Submitted (JUN-1997) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF008788; AAB63725.1; -.
DR   InterPro; IPR001364; Hemagglutn.
DR   Pfam; PF00509; Hemagglutinin; 1.
DR   ProDom; PD000225; Hemagglutn; 1.
FT   NON_TER       1      1
FT   NON_TER     326    326
SQ   SEQUENCE   326 AA;  36349 MW;  6A11A4CBBFBD6CC9 CRC64;


  Query Match             100.0%;  Score 55;  DB 12;  Length 326;
  Best Local Similarity   100.0%;  Pred. No. 0.074;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 YPYDVPDYA 9
          |||||||||
Db     98 YPYDVPDYA 106


RESULT  11
O92324
ID   O92324      PRELIMINARY;      PRT;   327 AA.
AC   O92324;
DT   01-NOV-1998 (TrEMBLrel. 08, Created)
DT   01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   HEMAGGLUTININ (FRAGMENT).
GN   HA.
OS   Influenza A virus.
OC   Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;
OC   Influenza virus A and B group; Influenza A viruses.
OX   NCBI_TaxID=11320;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=A/SYDNEY/05/97-LIKE(H3N2);
RA   Osiowy C.K.;
RT   "Genetic characterization of A/Sydney/05/97-like(H3N2) Influenza virus
RT   isolates circulating in Canada during the 1997/98 Influenza season.";
RL   Submitted (AUG-1998) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF087700; AAC36729.1; -.
DR   InterPro; IPR001364; Hemagglutn.
DR   Pfam; PF00509; Hemagglutinin; 1.
DR   ProDom; PD000225; Hemagglutn; 1.
FT   NON_TER       1      1
FT   NON_TER     327    327
SQ   SEQUENCE   327 AA;  36389 MW;  909239953649069E CRC64;


  Query Match             100.0%;  Score 55;  DB 12;  Length 327;
  Best Local Similarity   100.0%;  Pred. No. 0.074;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 YPYDVPDYA 9
          |||||||||
Db     98 YPYDVPDYA 106


RESULT  12
O92325
ID   O92325      PRELIMINARY;      PRT;   327 AA.
AC   O92325;
DT   01-NOV-1998 (TrEMBLrel. 08, Created)
DT   01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   HEMAGGLUTININ (FRAGMENT).
GN   HA.
OS   Influenza A virus.
OC   Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;
OC   Influenza virus A and B group; Influenza A viruses.
OX   NCBI_TaxID=11320;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=A/SYDNEY/05/97-LIKE(H3N2);
RA   Osiowy C.K.;
RT   "Genetic characterization of A/Sydney/05/97-like(H3N2) Influenza virus
RT   isolates circulating in Canada during the 1997/98 Influenza season.";
RL   Submitted (AUG-1998) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF087703; AAC36732.1; -.
DR   InterPro; IPR001364; Hemagglutn.
DR   Pfam; PF00509; Hemagglutinin; 1.
DR   ProDom; PD000225; Hemagglutn; 1.
FT   NON_TER       1      1
FT   NON_TER     327    327
SQ   SEQUENCE   327 AA;  36457 MW;  D741E3C39D0C4D6A CRC64;


  Query Match             100.0%;  Score 55;  DB 12;  Length 327;
  Best Local Similarity   100.0%;  Pred. No. 0.074;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 YPYDVPDYA 9
          |||||||||
Db     98 YPYDVPDYA 106


RESULT  13
O92327
ID   O92327      PRELIMINARY;      PRT;   327 AA.
AC   O92327;
DT   01-NOV-1998 (TrEMBLrel. 08, Created)
DT   01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   HEMAGGLUTININ (FRAGMENT).
GN   HA.
OS   Influenza A virus.
OC   Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;
OC   Influenza virus A and B group; Influenza A viruses.
OX   NCBI_TaxID=11320;
```

```
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=A/SYDNEY/05/97(H3N2);
RA   Osiowy C.K.;
RT   "Genetic characterization of A/Sydney/05/97-like(H3N2) Influenza virus
RT   isolates circulating in Canada during the 1997/98 Influenza season.";
RL   Submitted (AUG-1998) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF087709; AAC36738.1; -.
DR   InterPro; IPR001364; Hemagglutn.
DR   Pfam; PF00509; Hemagglutinin; 1.
DR   ProDom; PD000225; Hemagglutn; 1.
FT   NON_TER       1      1
FT   NON_TER     327    327
SQ   SEQUENCE   327 AA;  36293 MW;  4A72007829B091F5 CRC64;


  Query Match             100.0%;  Score 55;  DB 12;  Length 327;
  Best Local Similarity   100.0%;  Pred. No. 0.074;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 YPYDVPDYA 9
          |||||||||
Db     98 YPYDVPDYA 106


RESULT  14
Q9YT90
ID   Q9YT90      PRELIMINARY;      PRT;   327 AA.
AC   Q9YT90;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   HEMAGGLUTININ (FRAGMENT).
GN   HA.
OS   influenza A virus (A/Sydney/05/97-like(H3N2)).
OC   Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;
OC   Influenza virus A and B group; Influenza A viruses; Influenza A virus.
OX   NCBI_TaxID=82372;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=A/SYDNEY/05/97-LIKE(H3N2);
RA   Osiowy C.K.;
RT   "Genetic characterization of A/Sydney/05/97-like(H3N2) Influenza virus
RT   isolates circulating in Canada during the 1997/98 Influenza season.";
RL   Submitted (AUG-1998) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF087701; AAC36730.1; -.
DR   InterPro; IPR001364; Hemagglutn.
DR   Pfam; PF00509; Hemagglutinin; 1.
DR   PRINTS; PR00329; HEMAGGLUTN12.
DR   ProDom; PD000225; Hemagglutn; 1.
FT   NON_TER       1      1
FT   NON_TER     327    327
SQ   SEQUENCE   327 AA;  36373 MW;  ADBD67D856EBCC96 CRC64;


  Query Match             100.0%;  Score 55;  DB 12;  Length 327;
  Best Local Similarity   100.0%;  Pred. No. 0.074;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 YPYDVPDYA 9
          |||||||||
Db     98 YPYDVPDYA 106


RESULT  15
Q9YT89
ID   Q9YT89      PRELIMINARY;      PRT;   327 AA.
AC   Q9YT89;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   HEMAGGLUTININ (FRAGMENT).
GN   HA.
OS   influenza A virus (A/Sydney/05/97-like(H3N2)).
OC   Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;
OC   Influenza virus A and B group; Influenza A viruses; Influenza A virus.
OX   NCBI_TaxID=82372;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=A/SYDNEY/05/97-LIKE(H3N2);
RA   Osiowy C.K.;
RT   "Genetic characterization of A/Sydney/05/97-like(H3N2) Influenza virus
RT   isolates circulating in Canada during the 1997/98 Influenza season.";
RL   Submitted (AUG-1998) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF087702; AAC36731.1; -.
DR   InterPro; IPR001364; Hemagglutn.
DR   Pfam; PF00509; Hemagglutinin; 1.
DR   PRINTS; PR00329; HEMAGGLUTN12.
DR   ProDom; PD000225; Hemagglutn; 1.
FT   NON_TER       1      1
FT   NON_TER     327    327
SQ   SEQUENCE   327 AA;  36446 MW;  9E9FA8433CFC1B01 CRC64;


  Query Match             100.0%;  Score 55;  DB 12;  Length 327;
  Best Local Similarity   100.0%;  Pred. No. 0.074;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 YPYDVPDYA 9
          |||||||||
Db     98 YPYDVPDYA 106


Search completed: April 11, 2002, 10:22:57
Job time: 427 sec
```

GenCore version 4.5

OM protein - protein search, using sw model

Run on: April 11, 2002, 10:15:43 ; Search time 53.05 Seconds
(without alignments)
19.548 Million cell updates/sec

Title: US-09-284-787-2
Perfect score: 80
Sequence: 1 YPYDVPDYAGSGSK 14

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 522463 seqs, 74073290 residues

Total number of hits satisfying chosen parameters: 522463

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_1101:*
1: /SIDS8/gcgdata/geneseq/geneseqp/AA1980.DAT:*
2: /SIDS8/gcgdata/geneseq/geneseqp/AA1981.DAT:*
3: /SIDS8/gcgdata/geneseq/geneseqp/AA1982.DAT:*
4: /SIDS8/gcgdata/geneseq/geneseqp/AA1983.DAT:*
5: /SIDS8/gcgdata/geneseq/geneseqp/AA1984.DAT:*
6: /SIDS8/gcgdata/geneseq/geneseqp/AA1985.DAT:*
7: /SIDS8/gcgdata/geneseq/geneseqp/AA1986.DAT:*
8: /SIDS8/gcgdata/geneseq/geneseqp/AA1987.DAT:*
9: /SIDS8/gcgdata/geneseq/geneseqp/AA1988.DAT:*
10: /SIDS8/gcgdata/geneseq/geneseqp/AA1989.DAT:*
11: /SIDS8/gcgdata/geneseq/geneseqp/AA1990.DAT:*
12: /SIDS8/gcgdata/geneseq/geneseqp/AA1991.DAT:*
13: /SIDS8/gcgdata/geneseq/geneseqp/AA1992.DAT:*
14: /SIDS8/gcgdata/geneseq/geneseqp/AA1993.DAT:*
15: /SIDS8/gcgdata/geneseq/geneseqp/AA1994.DAT:*
16: /SIDS8/gcgdata/geneseq/geneseqp/AA1995.DAT:*
17: /SIDS8/gcgdata/geneseq/geneseqp/AA1996.DAT:*
18: /SIDS8/gcgdata/geneseq/geneseqp/AA1997.DAT:*
19: /SIDS8/gcgdata/geneseq/geneseqp/AA1998.DAT:*
20: /SIDS8/gcgdata/geneseq/geneseqp/AA1999.DAT:*
21: /SIDS8/gcgdata/geneseq/geneseqp/AA2000.DAT:*
22: /SIDS8/gcgdata/geneseq/geneseqp/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	80	100.0	14	19	AAW59451	Human influenza vi
2	75	93.8	165	22	AAG63605	Amino acid sequenc
3	71	88.8	34	21	AAB07730	Peptide encoded by
4	68	85.0	32	21	AAY96992	Peptide E16853A1 f
5	67	83.8	481	18	AAW27158	Saccharomyces cere
6	65	81.2	16	21	AAB03963	pTAT-HA linker pol
7	65	81.2	16	21	AAB29443	pTAT-HA linker-enc
8	65	81.2	17	18	AAW41032	Flu-epitope scramb
9	65	81.2	17	18	AAW29116	Flu epitope tagged
10	65	81.2	17	22	AAB45969	Transdominant effe
11	65	81.2	25	19	AAW48631	Chimeric adenovira
12	65	81.2	25	19	AAW46332	Haemagglutinin seq
13	63	78.8	15	17	AAR81452	Haemagglutinin tag
14	62	77.5	228	21	AAB10433	Expression vector
15	62	77.5	232	20	AAY00874	pHOOK-1.GM-CSF fus
16	62	77.5	352	20	AAY06272	Anti-Fc gamma rece
17	62	77.5	353	20	AAY06273	Anti Fc alpha rece
18	61	76.2	12	19	AAW77291	HA tag. Synthetic
19	61	76.2	353	16	AAR81355	Thrombopoietin pro
20	61	76.2	353	16	AAR81356	Thrombopoietin pro
21	61	76.2	371	20	AAW73369	Epitope tagged TBP
22	60.5	75.6	272	17	AAR92254	Neural cell adhesi
23	60	75.0	2272	18	AAW21731	GAL4/HA/NuMA fusio
24	59	73.8	14	17	AAW05097	HA1 epitope. Infl
25	59	73.8	18	20	AAW95558	Mus musculus haema
26	59	73.8	19	21	AAY90725	Influenza haemaggl
27	59	73.8	27	18	AAW08426	pSMTN3 fragment fo
28	59	73.8	27	20	AAY30054	Peptide encooded b
29	59	73.8	27	22	AAB69626	Plasma membrane ta
30	59	73.8	38	20	AAY49751	Compact structure
31	59	73.8	41	18	AAW08425	Construct pCGNN-ZF
32	59	73.8	181	21	AAB21192	Exo16 partial prot
33	59	73.8	359	20	AAW85003	Amino acid sequenc
34	59	73.8	359	21	AAY86206	LexA-GAL4 fusion p
35	59	73.8	389	21	AAY96709	HA-tagged I-kappa-
36	59	73.8	971	16	AAR76070	Yeast MSH1 protein
37	58	72.5	20	21	AAY96145	CSFV envelope prot
38	58	72.5	126	20	AAY32764	HM1.24 antigenic p
39	58	72.5	143	20	AAY32763	HM1.24 antigenic p
40	58	72.5	147	20	AAY32768	HM1.24 antigenic p
41	58	72.5	1070	18	AAW17789	Green fluorescent
42	57	71.2	13	19	AAW44013	Haemagglutinin pep
43	57	71.2	13	22	AAB30796	AMino acid sequenc
44	57	71.2	21	21	AAY49277	Sequence of angios
45	57	71.2	27	21	AAY49279	Sequence of angios

ALIGNMENTS

RESULT 1
AAW59451
ID AAW59451 standard; peptide; 14 AA.
XX
AC AAW59451;
XX
DT 28-AUG-1998 (first entry)
XX
DE Human influenza virus haemagglutinin epitope peptide fragment #2.
XX
KW Haemagglutinin; epitope; monoclonal antibody; detection; isolation;
KW fusion protein.
XX
OS Human parainfluenza virus.
XX
FH Key Location/Qualifiers
FT Modified-site 1
FT /note= "N-terminal acetyl group"
FT Modified-site 14
FT /note= "C-terminal epsilon-bionoyl-amide group"
XX
PN DE19643314-A1.
XX
PD 23-APR-1998.
XX
PF 21-OCT-1996; 96DE-1043314.
XX
PR 21-OCT-1996; 96DE-1043314.
XX
PA (BOEF) BOEHRINGER MANNHEIM GMBH.
XX
PI Emrich T, Grol M, Hinzpeter M;
XX
DR WPI; 1998-241603/22.

XX
PT Monoclonal antibodies to influenza virus haemagglutinin - useful for
PT detecting and isolating haemagglutinin proteins
XX
PS Claim 7; Page 2; 6pp; German.
XX
CC This peptide fragment represents a human influenza virus haemagglutinin
CC epitope which is used to make monoclonal antibodies with an affinity
CC constant of 108 M-1. Such antibodies are used for detecting and isolating
CC native human influenza virus haemagglutinin (HA), modified HA or HA
CC fusion proteins.
XX
SQ Sequence 14 AA;

Query Match 100.0%; Score 80; DB 19; Length 14;
Best Local Similarity 100.0%; Pred. No. 2.7e-06;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 YPYDVPDYAGSGSK 14
||||||||||||||
Db 1 ypydvpdyagsgsk 14

RESULT 2
AAG63605
ID AAG63605 standard; Protein; 165 AA.
XX
AC AAG63605;
XX
DT 15-OCT-2001 (first entry)
XX
DE Amino acid sequence of N-terminal fusion protein of pIX protein.
XX
KW pIX protein; chimeric protein; gene transfer vector.
XX
OS Synthetic.
OS Mastadenovirus.
XX
PN WO200158940-A2.
XX
PD 16-AUG-2001.
XX
PF 09-FEB-2001; 2001WO-US04233.
XX
PR 09-FEB-2000; 2000US-0181163.
XX
PA (GENV-) GENVEC INC.
XX
PI Roelvink PW, Kovesdi I, Wickham TJ;
XX
DR WPI; 2001-497066/54.
DR N-PSDB; AAH74658.
XX
PT Chimeric pIX protein useful in an adenovirus gene transfer vector for
PT infecting cells comprises at least one adenoviral pIX domain and a
PT non-native amino acid -
XX
PS Example 1; Page 26-27; 28pp; English.
XX
CC The present sequence represents a chimeric fusion protein of the
CC invention. The specification describes a chimeric pIX protein having
CC at least one adenoviral pIX domain and a non-native amino acid sequence.
CC The non-native sequence is a ligand that binds to a substrate present on
CC surface cells. The chimeric proteins are used for producing adenovirus
CC gene transfer vectors.
XX
SQ Sequence 165 AA;

Query Match 93.8%; Score 75; DB 22; Length 165;
Best Local Similarity 100.0%; Pred. No. 0.00022;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 YPYDVPDYAGSGS 13
|||||||||||||
Db 4 ypydvpdyagsgs 16

RESULT 3
AAB07730
ID AAB07730 standard; Peptide; 34 AA.
XX
AC AAB07730;
XX
DT 07-NOV-2000 (first entry)
XX
DE Peptide encoded by amplification product encoding a FLAG-HA peptide.
XX
KW Genetic element; optical property; gene product; microcapsule.
XX
OS Synthetic.
XX
PN WO200040712-A1.
XX
PD 13-JUL-2000.
XX
PF 06-JAN-2000; 2000WO-GB00030.
XX
PR 07-JAN-1999; 99GB-0000298.
XX
PA (MEDI-) MEDICAL RES COUNCIL.
XX
PI Griffiths A, Tawfik D;
XX
DR WPI; 2000-465978/40.
DR N-PSDB; AAA59266.
XX
PT Isolation of genetic elements encoding a gene product with a desired
PT activity is carried out in microcapsules where changes in optical
PT activity are detected -
XX
PS Example 12; Page 79-80; 128pp; English.
XX
CC The specification describes a method for isolating one or more genetic
CC elements encoding a gene product with a desired activity. Expression of
CC the genetic element results, directly or indirectly in the modification
CC of an optical property of the encoded polypeptide. The method comprises
CC compartmentalizing the genetic elements into microcapsules, expressing
CC them to produce their respective gene products within the microcapsules,
CC and sorting genetic elements which produce the gene product(s) with the
CC desired activity. The method is used for isolating genetic elements
CC which encode a gene product with a desired activity. This can be a
CC catalytic, regulatory or binding activity. The present sequence
CC represents a fusion of a flag epitope and influenza haemagglutinin. It
CC was used in the course of the invention.
XX
SQ Sequence 34 AA;

Query Match 88.8%; Score 71; DB 21; Length 34;
Best Local Similarity 92.3%; Pred. No. 0.00018;
Matches 12; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 YPYDVPDYAGSGS 13
||||||||| ||
Db 22 ypydvpdyagggs 34

RESULT 4
AAY96992
ID AAY96992 standard; peptide; 32 AA.
XX
AC AAY96992;
XX
DT 31-OCT-2000 (first entry)

XX
DE Peptide E16853A1 for construction of rTether1 containing HA-tag.
XX
KW PTH; parathyroid hormone; conjugate; bone mass; bone reformation;
KW PTH-1 receptor; resorption; remodeling; tether1; osteoporosis.
XX
OS Synthetic.
XX
PN WO200039278-A2.
XX
PD 06-JUL-2000.
XX
PF 30-DEC-1999; 99WO-US31108.
XX
PR 31-DEC-1998; 98US-0114577.
XX
PA (GARD/) GARDELLA T J.
PA (KRON/) KRONENBERG H M.
PA (POTT/) POTTS J T.
PA (JUEP/) JUEPPNER H.
XX
PI Gardella TJ, Kronenberg HM, Potts JT, Jueppner H;
XX
DR WPI; 2000-452384/39.
DR N-PSDB; AAA51743.
XX
PT New compound comprising an amino terminal signaling functional domain
PT linked to a carboxy-terminal binding portion of parathyroid hormone for
PT treating mammalian conditions characterized by decreases in bone mass
XX
PS Example 4; Fig 11G; 119pp; English.
XX
CC AAA51738-44 were used to construct a chimeric rat PTH-1 receptor,
CC rTether1 and control plasmids.
CC Compounds of the structure or formula S-(L)_n-B, R_1-S-(L)_n-R or
CC S-(L)_n-R, are new. S is an amino terminal signaling functional domain
CC of parathyroid hormone (PTH); L is a linker molecule present n times
CC (where n is a positive integer 1-10 (preferably 4)); B is a C-terminal
CC binding portion of PTH(1-34) or a PTH-related protein (PTH-rP) (1-34);
CC R_1 is the PTH-1 receptor signal sequence; and R is (a portion of) PTH-1
CC receptor sequence. The new compounds are used for treating mammalian
CC conditions characterized by decreases in bone mass, determining rates of
CC bone reformation, bone resorption and/or bone remodeling, treating
CC diseases and disorders associated with decreased tether1 activity,
CC increasing cAMP in a mammalian cell having PTH-1 receptors, or screening
CC for a peptide or non-peptide PTH (claimed). The new compound can be
CC administered by inhalation unlike the large native PTH or PTHrP which
CC avoids the need for regular injections to treat osteoporosis.
XX
SQ Sequence 32 AA;

Query Match 85.0%; Score 68; DB 21; Length 32;
Best Local Similarity 78.6%; Pred. No. 0.00049;
Matches 11; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

```
Qy       1 YPYDVPDYAGSGSK 14
           |||||||||| | :
Db      10 ypydvpdyagggge 23
```

RESULT 5
AAW27158
ID AAW27158 standard; Protein; 481 AA.
XX
AC AAW27158;
XX
DT 20-APR-1998 (first entry)
XX
DE Saccharomyces cerevisiae alpha mating factor receptor Ste2.
XX
KW Alpha mating factor receptor; Ste2; mutant; biosensor; diagnosis;
KW G-protein coupled receptor; GPCR; environmental monitoring;
KW on-line process monitoring.
XX
OS Saccharomyces cerevisiae.
XX
PN WO9735985-A1.
XX
PD 02-OCT-1997.
XX
PF 19-MAR-1997; 97WO-GB00746.
XX
PR 22-MAR-1996; 96GB-0006126.
XX
PA (BRBI-) BRITISH BIOTECH PHARM LTD.
XX
PI Edwards RM;
XX
DR WPI; 1997-489648/45.
DR N-PSDB; AAT85308.
XX
PT Cells for surface expression of mutants of natural G-protein
PT receptor - where the mutants can bind a ligand that the natural
PT receptor cannot, useful as biosensors
XX
PS Example 1; Pages 13-15; 35pp; English.
XX
CC The present sequence is the Saccharomyces cerevisiae alpha
CC mating factor receptor Ste2, which was used in the development of
CC a method to prepare cells presenting on their surface a mutant
CC G-protein coupled receptor (GPCR) binding a desired ligand which
CC the natural receptor cannot. The method comprises generating
CC replicable expression plasmids comprising a coding sequence for a
CC GPCR mutated by one or more amino acids in its transmembrane and/or
CC extracellular domains, and operatively linked expression control
CC sequences, transforming host cells with the plasmids and contacting
CC the cells with the ligand and isolating those which bind. GPCR are
CC natural receptor proteins, i.e. membrane proteins that allow cells
CC to detect molecules in the extracellular environment. Cells
CC presenting mutant GPCR on their surfaces are useful as biosensors,
CC with potential applications in diagnostics and environmental and
CC on-line process monitoring.
XX
SQ Sequence 481 AA;

Query Match 83.8%; Score 67; DB 18; Length 481;
Best Local Similarity 92.3%; Pred. No. 0.012;
Matches 12; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy       1 YPYDVPDYAGSGS 13
           |||||||||||  |
Db     447 ypydvpdyagsts 459
```

RESULT 6
AAB03963
ID AAB03963 standard; Peptide; 16 AA.
XX
AC AAB03963;
XX
DT 26-FEB-2001 (first entry)
XX
DE pTAT-HA linker polypeptide.
XX
KW Chimeric protein; fusion protein; FLICE like inhibitor protein;
KW FLIP; Fas; TNF; apoptosis; caspase-8; ligand; T cell; thymocyte;
KW tumour specific antigen; immune response; therapy; prophylaxis;
KW diagnosis; HIV; human immunodeficiency syndrome; AIDS;
KW acquired immune deficiency syndrome.
XX
OS Human immunodeficiency virus.
XX
PN WO200059935-A1.
XX

PD 12-OCT-2000.
XX
PF 05-APR-2000; 2000WO-US09002.
XX
PR 05-APR-1999; 99US-0127867.
PR 06-APR-1999; 99US-0128021.
XX
PA (MAYO-) MAYO FOUND MEDICAL EDUCATION & RES.
PA (PAYA/) PAYA C.
PA (ALGE/) ALGECIRAS-SCHMINICH A.
XX
PI Paya C, Algeciras-schminich A;
XX
DR WPI; 2000-664988/64.
DR N-PSDB; AAA54300.
XX
PT Fusion polypeptide useful for inhibiting ligand-induced apoptosis,
PT comprises portion of anti-apoptotic polypeptide linked to a transport
PT group
XX
PS Disclosure; Fig 9; 89pp; English.
XX
CC A chimeric group or fusion peptide which comprises a portion of an
CC anti-apoptotic polypeptide which inhibits apoptosis of lymphocytes in
CC combination with a transport group is described. The transport group
CC is capable of transporting the chimeric group or fusion peptide
CC across the cell membrane. The anti-apoptotic polypeptide is FLICE-like
CC inhibitor protein (FLIP) which inhibits Fas and TNF mediated apoptosis
CC by inhibiting binding of Caspase-8 to the Fas receptor complex, thus
CC shutting off the downstream Fas signalling pathway. The chimeric group
CC and fusion peptide are useful for inhibiting ligand-induced apoptosis
CC by bringing them into contact with T cells. The chimeric group is
CC useful for expanding T cells in vitro e.g. T cells specific for
CC particular antigens such as tumour-specific antigen, for enhancing
CC immune response and to inhibit the apoptosis of chronically activated
CC T cells e.g. activated CD4^+ T cells in HIV infected patients. The
CC chimeric group is also useful for therapeutic, prophylactic or
CC diagnosis of intracellular delivery of small molecules and
CC macromolecules such as anti-apoptotic polypeptides and nucleic
CC acids encoding such polypeptides. Two primers (AAA54297, AAA54298) were
CC used to amplify the FLIP cDNA for subcloning into the XhoI-NcoI
CC site of the pHA-TAT vector which contains the N-terminal protein
CC transduction domain from the human immunodeficiency virus tat
CC protein. Tat is a preferred transport moiety.
XX
SQ Sequence 16 AA;

Query Match 81.2%; Score 65; DB 21; Length 16;
Best Local Similarity 100.0%; Pred. No. 0.00071;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 YPYDVPDYAGS 11
||||||||||||
Db 4 ypydvpdyags 14

RESULT 7
AAB29443
ID AAB29443 standard; peptide; 16 AA.
XX
AC AAB29443;
XX
DT 09-FEB-2001 (first entry)
XX
DE pTAT-HA linker-encoded peptide.
XX
KW Protein transduction domain; fusion molecule; therapeutic agent;
KW drug targetting; drug discovery; cell transduction; bioavailability;
KW vaccine; nervous system disorder; Alzheimer's disease;
KW Parkinson's disease; Huntington's disease; pre-senile dementia; epilepsy;
KW seizure; compulsive behaviour; meningitis; encephalitis; ischaemia;
KW spongiform encephalopathy; dyslexia; age-related memory loss;
KW Lou Gehring's disease; viral infection; HIV; bacterial infection.
XX
OS Human immunodeficiency virus.
OS Synthetic.
XX
PN WO200062067-A1.
XX
PD 19-OCT-2000.
XX
PF 28-FEB-2000; 2000WO-US05097.
XX
PR 28-FEB-1999; 99US-0122757.
PR 29-AUG-1999; 99US-0151291.
XX
PA (UNIW) UNIV WASHINGTON.
XX
PI Dowdy SF;
XX
DR WPI; 2000-647439/62.
DR N-PSDB; AAC63859.
XX
PT Fusion molecules comprising protein transduction domains and
PT therapeutic agents, useful for treating e.g. Alzheimer's and
PT Parkinson's diseases, dementia and epilepsy -
XX
PS Example 1; Fig 2; 191pp; English.
XX
CC The invention relates to a novel fusion molecule comprising at least
CC one protein transduction domain (PTD) and at least one linked molecule,
CC where the linked molecule has therapeutic or prophylactic activity
CC against a medical condition. The invention also relates to methods of
CC drug discovery in which the test compound is linked to a suitable
CC transducing protein and introduced to a cell; a method of killing
CC resistant microorganisms using a suitable fusion molecule; a mammal
CC comprising a covalently linked fusion molecule; and a mammal adapted for
CC experimental use in which at least one transduction molecule has been
CC transduced into essentially all the cells of the mammal. The fusion
CC molecule is used to deliver a therapeutic agent to a mammal, especially
CC a human. The linked molecule may be a vaccine, an anti-infective drug,
CC a cardiovascular drug, an antitumour drug, an analgesic, an
CC antiinflammatory, a diagnostic marker or a drug for the treatment or
CC prevention of a central or peripheral nervous system disorder. The
CC central nervous system (CNS) disorder is especially Alzheimer's disease,
CC Parkinson's disease, Huntington's disease, and also includes pre-senile
CC dementia, epilepsy and seizures, compulsive behaviour, meningitis
CC (including viral and bacterial meningitis), encephalitis, ischaemia,
CC scrapie (or related spongiform encephalopathies), dyslexia, age-related
CC memory loss or Lou Gehring's disease. Fusion molecules can also be
CC used to kill virally infected cells, especially those infected with HIV.
CC The vaccines are used to treat or prevent bacterial or viral infections.
CC The methods are a highly effective means for transducing a molecule
CC into an entire mammal or into specific cells, tissues, organs and
CC systems within it. They also overcome bioavailability problems that
CC are associated with many therapeutic agents (e.g., large molecular size,
CC hydrophobicity, hydrophilicity, biological resistance), by providing
CC efficient transduction of the target cell. The present sequence
CC represents a protein transduction domain used in the invention.
XX
SQ Sequence 16 AA;

Query Match 81.2%; Score 65; DB 21; Length 16;
Best Local Similarity 100.0%; Pred. No. 0.00071;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 YPYDVPDYAGS 11
||||||||||||
Db 4 ypydvpdyags 14

RESULT 8
AAW41032
ID AAW41032 standard; Protein; 17 AA.

XX
AC AAW41032;
XX
DT 12-MAY-1998 (first entry)
XX
DE Flu-epitope scrambled peptide.
XX
KW Transdominant bioactive agent; phenotype alteration; signalling pathway;
KW antitumour agent; cardiovascular disease; angiogenesis; atherosclerosis;
KW obesity; neurodegeneration; bone disease; infection; allergy; therapy;
KW flu-epitope peptide.
XX
OS Synthetic.
OS Influenza virus.
XX
PN WO9727213-A1.
XX
PD 31-JUL-1997.
XX
PF 23-JAN-1997; 97WO-US01048.
XX
PR 23-JAN-1996; 96US-0589911.
PR 23-JAN-1996; 96US-0589109.
XX
PA (STRD) UNIV LELAND STANFORD JUNIOR.
XX
PI Noaln GP, Rothenberg SM;
XX
DR WPI; 1997-393612/36.
XX
PT Screening for trans-dominant active agents able to alter cell
PT phenotype - useful to identify potential drugs with e. g.
PT anti-tumour activity
XX
PS Example 1; Page 53; 91pp; English.
XX
CC This sequence represents a scrambled flu-epitope peptide, and can be
CC used in the method of the invention. The method is for screening for
CC transdominant bioactive agents (A) able to alter the phenotype of a cell
CC comprises: (a) introducing a library of different randomised nucleic
CC acids (I) into cells; and (b) screening the cells for altered phenotype
CC due to presence of (A). (A) are potential pharmaceuticals, from their
CC effect on signalling pathways. A particular application is to identify
CC peptides, encoded by (I), that have antitumour activity or are able to
CC increase sensitivity to, or reduce toxicity of, known antitumour agents.
CC More generally the method can identify agents for treatment of many other
CC conditions, e.g. cardiovascular diseases, angiogenesis, atherosclerosis,
CC obesity, neurodegeneration, bone disease, viral and other infections,
CC allergy etc.
XX
SQ Sequence 17 AA;

Query Match 81.2%; Score 65; DB 18; Length 17;
Best Local Similarity 100.0%; Pred. No. 0.00076;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 YPYDVPDYAGS 11
||||||||||||
Db 5 ypydvpdyags 15

RESULT 9
AAW29116
ID AAW29116 standard; Peptide; 17 AA.
XX
AC AAW29116;
XX
DT 04-FEB-1998 (first entry)
XX
DE Flu epitope tagged scrambled peptide.
XX
KW Transdominant intracellular effector peptide; RNA; screening;
KW altered phenotype; retrovirus; library; signal transduction;
KW antitumour; pBABE; influenza virus; flu; epitope.
XX
OS Synthetic.
XX
PN WO9727212-A1.
XX
PD 31-JUL-1997.
XX
PF 23-JAN-1997; 97WO-US01019.
XX
PR 23-JAN-1996; 96US-0589911.
PR 23-JAN-1996; 96US-0589109.
XX
PA (RIGE-) RIGEL PHARM INC.
XX
PI Noaln GP;
XX
DR WPI; 1997-393611/36.
XX
PT Screening for transdominant intracellular active agents able to
PT alter cell phenotype - useful for examining for changed phenotype,
PT particularly to identify potential drugs with e. g. antitumour
PT activity
XX
PS Example 1; Page 53; 91pp; English.
XX
CC This peptide comprises a Flu epitope tagged scrambled peptide.
CC Murine 10T1/2 Clone 8 cells were infected with pBabe puro
CC retrovirus constructs containing a sequence coding for the Flu
CC scrambled epitope or an inhibitor peptide (see AAW29117) identical
CC to the C2 region of protein kinase C (PKA) isozymes. Control clone
CC 8 cells showed predominantly cytoplasmic and perinuclear staining,
CC while PMA induced cells showed translocation to the nucleus. Cells
CC infected with constructs coding for the scrambled peptide showed
CC similar staining. Those infected with constructs coding for
CC the C2 region showed predominantly cytoplasmic and perinuclear
CC staining in both control and PMA induced cells. Novel outcomes can
CC occur upon expression of peptides in cells. A claimed method of
CC screening for transdominant intracellular bioactive agents (A) able
CC to alter the phenotype of a cell comprises: (a) introducing a
CC library of different randomised nucleic acids (I) into cells; and
CC (b) screening the cells for altered phenotype due to presence of
CC (A). Also new are: (1) a library of retroviruses containing
CC different randomised (I); and (2) a library of mammalian cells
CC containing the library of (1), preferably integrated into the
CC genome.
XX
SQ Sequence 17 AA;

Query Match 81.2%; Score 65; DB 18; Length 17;
Best Local Similarity 100.0%; Pred. No. 0.00076;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 YPYDVPDYAGS 11
||||||||||||
Db 5 ypydvpdyags 15

RESULT 10
AAB45969
ID AAB45969 standard; Peptide; 17 AA.
XX
AC AAB45969;
XX
DT 02-APR-2001 (first entry)
XX
DE Transdominant effector peptide associated screening peptide #49.
XX
KW Intracellular transdominant bioactive agent; screening; cell phenotype;
KW effector peptide.
XX

OS Unidentified.
XX
PN US6153380-A.
XX
PD 28-NOV-2000.
XX
PF 23-JAN-1997; 97US-0789333.
XX
PR 23-JAN-1996; 96US-0589108.
PR 23-JAN-1996; 96US-0589911.
XX
PA (STRD) UNIV LELAND STANFORD JUNIOR.
PA (RIGE-) RIGEL PHARM INC.
XX
PI Rothenberg SM, Nolan GP;
XX
DR WPI; 2001-060084/07.
XX
PT Methods for screening intracellular transdominant effector peptides and
PT RNA molecules comprise delivering random oligonucleotides to cells,
PT which are then screened for an altered phenotype -
XX
PS Example 1; Column 79-80; 57pp; English.
XX
CC This invention describes novel in vitro screening methods (I) for a
CC transdominant intracellular bioactive agent capable of altering the
CC phenotype of a cell. (I) comprises: (a) introducing a molecular library
CC of randomized candidate nucleic acids into several cells; and (b)
CC screening the cells for a cell exhibiting an altered phenotype, where the
CC altered phenotype is due to the presence of a transdominant bioactive
CC agent. The methods are particularly useful for screening intracellular
CC transdominant effector peptides and RNA molecules selected inside living
CC cells from randomized pools. (I) is also useful for introducing random
CC libraries into cells to screen for bioactive compounds. The methods allow
CC rapid and highly efficient screening of large numbers of random
CC oligonucleotides and their corresponding expression products in a single
CC step. In addition, the methods allow screening in the absence of
CC significant prior characterization of the cellular defect.
XX
SQ Sequence 17 AA;

Query Match 81.2%; Score 65; DB 22; Length 17;
Best Local Similarity 100.0%; Pred. No. 0.00076;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 YPYDVPDYAGS 11
||||||||||||
Db 5 ypydvpdyags 15

RESULT 11
AAW48631
ID AAW48631 standard; peptide; 25 AA.
XX
AC AAW48631;
XX
DT 11-SEP-1998 (first entry)
XX
DE Chimeric adenoviral penton base protein epitope 3.
XX
KW Short-shafted adenoviral fibre; ssAF; penton base; cystic fibrosis;
KW bispecific antibody; HIV; cancer.
XX
OS Unidentified.
XX
PN WO9807877-A1.
XX
PD 26-FEB-1998.
XX
PF 21-AUG-1997; 97WO-US14718.
XX
PR 21-AUG-1996; 96US-0700846.
XX
PA (GENV-) GENVEC INC.
XX
PI Kovesdi I, Roelvink PW, Wickham TJ;
XX
DR WPI; 1998-169180/15.
XX
PT Short-shafted adenoviral fibre constructs - used for producing
PT products that facilitate entry into target cells, useful for gene
PT therapy of cellular disorders
XX
PS Disclosure; Page 39; 98pp; English.
XX
CC The present peptide sequence is incorporated into an adenoviral fiber.
CC The fiber protein and its penton base are used by the adenovirus to
CC interact with distinct cellular receptors to attach to and efficiently
CC infect a cell. The invention claims for a transfer vector which contains
CC DNA encoding a short-shafted adenoviral fibre (ssAF). The adenoviral
CC penton base can be modified by incorporating sequences for a ligand to a
CC cell surface receptor or sequences that allow binding to a bispecific
CC antibody. Bispecific antibodies raised against the present peptide
CC allows one end of the antibody molecule to bind to the fiber and the
CC other end to bind specifically to a cell surface receptor. Therefore
CC this method enables targeting of the adenovirus to a desired cell-surface
CC receptor by the introduction of a non-native amino acid sequence either
CC into the penton base or the fibre knob. The non-native amino acid
CC sequence can be such that it enables direct or indirect binding, e.g. by
CC means of a bispecific or multispecific binding agent, of the adenovirus
CC to the desired cell-surface receptor. Using the a short shafted
CC adenovirus to attach to a cell and therefore effect entry into the cell
CC would reduce the level or efficiency of adenoviral fibre binding to its
CC cell-surface receptor and increase the adenoviral penton base binding to
CC its cell-surface receptor, thereby increasing the specificity of binding
CC of the adenovirus to a given cell. The transfer vector is claimed
CC to be useful for expressing therapeutic genes, for treating inherited
CC diseases such as cystic fibrosis or for inducing cell death in
CC pathogenic infections (e.g. HIV infections) and cancers.
XX
SQ Sequence 25 AA;

Query Match 81.2%; Score 65; DB 19; Length 25;
Best Local Similarity 100.0%; Pred. No. 0.0011;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 YPYDVPDYAGS 11
||||||||||||
Db 14 ypydvpdyags 24

RESULT 12
AAW46332
ID AAW46332 standard; Peptide; 25 AA.
XX
AC AAW46332;
XX
DT 08-MAY-1998 (first entry)
XX
DE Haemagglutinin sequence of a chimeric adenovirus penton base protein.
XX
KW Integrin; cell surface receptor; penton base protein; adenovirus;
KW binding site; binding domain; cell surface binding site; gene therapy;
KW bispecific molecule; adenoviral transfer vector; pAT.
XX
OS Synthetic.
XX
PN US5712136-A.
XX
PD 27-JAN-1998.
XX
PF 17-APR-1996; 96US-0634060.
XX
PR 08-SEP-1994; 94US-0303162.

XX
PA (GENV-) GENVEC INC.
XX
PI Brough DE, Bruder JT, Kovesdi I, McVey DL, Roelvink PW;
PI Wickham TJ;
XX
DR WPI; 1998-119984/11.
DR N-PSDB; AAV04953-54.
XX
PT Methods for introducing adenovirus into cells - used for genetic
PT engineering and gene therapy
XX
PS Example 13; Column 34; 56pp; English.
XX
CC The present sequence represents a haemagglutinin sequence of a chimeric
CC adenovirus penton base protein. The peptide is encoded by overlapping
CC oligonucleotides. The penton base protein binds to cell surface receptors
CC called integrins. The integrins not only provide a binding site for the
CC adenoviral penton base protein, but also mediate cellular adhesion to
CC the extracellular matrix molecules. The specification describes a method
CC of introducing an adenovirus into a cell in vitro having a particular
CC cell surface binding site. The adenovirus is contacted with a bispecific
CC molecule comprising a component that selectively binds a binding domain
CC of the penton base protein of the adenovirus and a second component that
CC selectively binds the cell surface binding site. A complex of the
CC adenovirus and the bispecific molecule is formed, and the cell is
CC contacted with it to allow entry of the adenovirus into the cell. The
CC methods can be used for research and the vectors can be used for gene
CC therapy.
XX
SQ Sequence 25 AA;

Query Match 81.2%; Score 65; DB 19; Length 25;
Best Local Similarity 100.0%; Pred. No. 0.0011;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 YPYDVPDYAGS 11
||||||||||||
Db 14 ypydvpdyags 24

RESULT 13
AAR81452
ID AAR81452 standard; Peptide; 15 AA.
XX
AC AAR81452;
XX
DT 02-JUL-1996 (first entry)
XX
DE Haemagglutinin tag peptide.
XX
KW Bcl-Y; apoptosis; cell proliferation; cell death; diagnosis;
KW gene therapy; haemagglutinin; HA tag.
XX
OS Synthetic.
XX
PN WO9605232-A1.
XX
PD 22-FEB-1996.
XX
PF 09-AUG-1995; 95WO-US10103.
XX
PR 11-OCT-1994; 94US-0321071.
PR 09-AUG-1994; 94US-0287427.
XX
PA (IMMU-) IMMUNOGEN INC.
XX
PI Chittenden TD;
XX
DR WPI; 1996-139648/14.
XX
PT New isolated human Bcl-Y protein - used to develop prods. for
PT treating disorders characterised by inappropriate cell proliferation
PT or cell death
XX
PS Example 1; Page 65; 100pp; English.
XX
CC A haemagglutinin (HA) tag peptide (AAR81452) was expressed as a
CC fusion with amino acids 2-221 of apoptosis-associated protein
CC Bcl-Y (see AAR81451) following PCR amplification (see AAT17378-79)
CC of Bcl-Y cDNA and insertion into a CMV-based pcDNA1/amp
CC vector derivative. The HA-Bcl-Y protein was produced in vitro
CC and in transfected COS7 and NIH3T3 cells.
XX
SQ Sequence 15 AA;

Query Match 78.8%; Score 63; DB 17; Length 15;
Best Local Similarity 84.6%; Pred. No. 0.0014;
Matches 11; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 YPYDVPDYAGSGS 13
|||||||||| ||
Db 3 ypydvpdyaslgs 15

RESULT 14
AAB10433
ID AAB10433 standard; Protein; 228 AA.
XX
AC AAB10433;
XX
DT 01-DEC-2000 (first entry)
XX
DE Expression vector pSEX11G2 protein G.
XX
KW Expression vector; antibody binding protein; monoclonal antibody; Neo-R;
KW B lymphocyte; myeloma cell; hybridoma; protein G; bla protein.
XX
OS Synthetic.
XX
PN DE19900635-A1.
XX
PD 13-JUL-2000.
XX
PF 11-JAN-1999; 99DE-1000635.
XX
PR 11-JAN-1999; 99DE-1000635.
XX
PA (DEKR-) DEUT KREBSFORSCHUNGSZENTRUM.
XX
PI Breitling F, Poustka A, Moldenhauer G;
XX
DR WPI; 2000-499832/45.
DR N-PSDB; AAA71429.
XX
PT Selecting monoclonal antibodies, by expressing them on the surface of
PT hybridomas attached to antibody-binding protein, then reaction with
PT antibody library -
XX
PS Claim 16; Fig 2; 22pp; German.
XX
CC This invention describes a novel method for the selection of monoclonal
CC antibodies (MAb) which comprises (i) fusing B lymphocytes with myeloma
CC cells to produce antibody-producing hybridomas such that the antibodies
CC are presented at the surface of the hybridomas by an antibody-binding
CC protein (I); and (ii) binding the antibody to antigens (Ag). The
CC invention also describes antibody-binding proteins (I) that comprise a
CC combination of the signal peptide of a murine immunoglobulin (Ig) kappa
CC chain or a murine MHC (major histocompatibility complex) Class I k(k)
CC molecule; an antibody-binding site of proteins A, G, L or LG, and the
CC transmembrane domain of PDGFR (platelet-derived growth factor receptor)
CC or CD52. The method is used to select MAb with specificity for particular
CC antigens. MAb can be selected without separate culture of hybridomas,
CC and selection can be made against many antigens in a library, optionally

CC on the basis of strength of affinity for a particular antigen. Complex
CC mixtures of hybridomas can be used for selection, reducing the time and
CC cost involved in MAb selection. This sequence represents the protein G
CC contained in the expression vector pSEX11G2 which contains the protein G,
CC Neo-R and the bla protein described in the method of the invention.
XX
SQ Sequence 228 AA;

Query Match 77.5%; Score 62; DB 21; Length 228;
Best Local Similarity 90.9%; Pred. No. 0.034;
Matches 10; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 YPYDVPDYAGS 11
||||||||||:
Db 22 ypydvpdyaga 32

RESULT 15
AAY00874
ID AAY00874 standard; Protein; 232 AA.
XX
AC AAY00874;
XX
DT 21-MAY-1999 (first entry)
XX
DE pHOOK-1.GM-CSF fusion protein sequence.
XX
KW GM-CSF; granulocyte macrophage colony stimulating factor; vaccine;
KW membrane-bound fusion protein; non-antibody immunomodulator; infection;
KW membrane attachment domain; cancer; autoimmune disease; therapy.
XX
OS Chimeric - Mus sp.
OS Chimeric - Synthetic.
XX
FH Key Location/Qualifiers
FT Misc-difference 220
FT /note= "unspecified amino acid"
XX
PN WO9906544-A1.
XX
PD 11-FEB-1999.
XX
PF 28-JUL-1998; 98WO-US15622.
XX
PR 29-JUL-1997; 97US-0902516.
XX
PA (IMMU-) IMMUNE RESPONSE CORP.
XX
PI Soo Hoo W;
XX
DR WPI; 1999-153774/13.
DR N-PSDB; AAX27068.
XX
PT Vaccine having a membrane-bound fusion protein - comprising an
PT immunomodulator and heterologous membrane attachment domain useful
PT for modulation of immune response against a disease-associated
PT antigen
XX
PS Example 1; Fig 2a-b; 91pp; English.
XX
CC This sequence is the pHOOK-1.GM-CSF fusion protein. The
CC invention relates to a vaccine comprising a cell having a membrane-bound
CC fusion protein (FP) consisting of a non-antibody immunomodulator (I)
CC linked to a heterologous membrane attachment domain. The vaccines are
CC used to treat or prevent a wide variety of cancers (e.g. of colon, breast
CC or prostate, melanoma, glioma), alone or as adjunct to other therapies;
CC autoimmune diseases (e.g. rheumatoid arthritis, psoriasis, multiple
CC sclerosis, systemic lupus erythematosus, type I diabetes, allergy)
CC or viral, bacterial or parasitic infections (e.g. human immune
CC deficiency virus, Helicobacter pylori, Porphyromonas gingivalis, or
CC Candida albicans). A panel of cells (or cell lines) each genetically
CC modified to express different antigens can be maintained as a repository,
CC for treatment or prevention of the appropriate tumour type, e.g. based on
CC histological analysis. Membrane-bound (I) provides a vaccine with
CC increased receptor-cytokine avidity, and thus stimulates a stronger
CC immune response.
XX
SQ Sequence 232 AA;

Query Match 77.5%; Score 62; DB 20; Length 232;
Best Local Similarity 90.9%; Pred. No. 0.035;
Matches 10; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 YPYDVPDYAGS 11
||||||||||:
Db 22 ypydvpdyaga 32

Search completed: April 11, 2002, 10:15:43
Job time: 93 sec

GenCore version 4.5

OM protein - protein search, using sw model

Run on: April 11, 2002, 10:14:43 ; Search time 24.73 Seconds
(without alignments)
12.739 Million cell updates/sec

Title: US-09-284-787-2
Perfect score: 80
Sequence: 1 YPYDVPDYAGSGSK 14

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 212252 seqs, 22503292 residues

Total number of hits satisfying chosen parameters: 212252

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	65	81.2	16	4	US-09-208-966-15	Sequence 15, Appl
2	65	81.2	17	4	US-08-789-333F-51	Sequence 51, Appl
3	65	81.2	25	1	US-08-634-060-27	Sequence 27, Appl
4	65	81.2	25	2	US-08-700-846-10	Sequence 10, Appl
5	63	78.8	15	1	US-08-321-071A-5	Sequence 5, Appli
6	62	77.5	219	2	US-08-902-516-2	Sequence 2, Appli
7	61	76.2	10	3	US-08-990-140-13	Sequence 13, Appl
8	60.5	75.6	102	1	US-08-282-951-2	Sequence 2, Appli
9	60.5	75.6	272	1	US-08-282-951-6	Sequence 6, Appli
10	59	73.8	14	2	US-08-723-726-1	Sequence 1, Appli
11	59	73.8	16	3	US-08-390-874C-13	Sequence 13, Appl
12	59	73.8	16	4	US-09-265-772-13	Sequence 13, Appl
13	59	73.8	19	4	US-09-169-015-52	Sequence 52, Appl
14	59	73.8	27	4	US-08-672-213-39	Sequence 39, Appl
15	59	73.8	40	4	US-08-672-213-29	Sequence 29, Appl
16	59	73.8	389	2	US-08-802-322-14	Sequence 14, Appl
17	57	71.2	13	1	US-08-660-626-11	Sequence 11, Appl
18	57	71.2	13	4	US-09-031-168-11	Sequence 11, Appl
19	55.5	69.4	32	2	US-09-066-074-7	Sequence 7, Appli
20	55.5	69.4	32	2	US-08-555-912A-7	Sequence 7, Appli
21	55	68.8	9	1	US-08-011-398B-6	Sequence 6, Appli
22	55	68.8	9	1	US-08-090-148-2	Sequence 2, Appli
23	55	68.8	9	1	US-08-370-225-6	Sequence 6, Appli
24	55	68.8	9	1	US-08-167-982-1	Sequence 1, Appli
25	55	68.8	9	1	US-08-634-060-31	Sequence 31, Appl
26	55	68.8	9	1	US-08-464-051-6	Sequence 6, Appli
27	55	68.8	9	1	US-08-461-859-6	Sequence 6, Appli
28	55	68.8	9	1	US-08-470-837-28	Sequence 28, Appl
29	55	68.8	9	2	US-08-478-386A-19	Sequence 19, Appl
30	55	68.8	9	2	US-08-292-597-19	Sequence 19, Appl
31	55	68.8	9	2	US-08-462-498-6	Sequence 6, Appli
32	55	68.8	9	2	US-08-922-267A-60	Sequence 60, Appl
33	55	68.8	9	2	US-08-388-653-19	Sequence 19, Appl
34	55	68.8	9	2	US-08-473-985-19	Sequence 19, Appl
35	55	68.8	9	2	US-08-690-011A-17	Sequence 17, Appl
36	55	68.8	9	2	US-08-690-011A-49	Sequence 49, Appl
37	55	68.8	9	2	US-08-685-625A-4	Sequence 4, Appli
38	55	68.8	9	2	US-08-687-559-14	Sequence 14, Appl
39	55	68.8	9	2	US-08-700-846-11	Sequence 11, Appl
40	55	68.8	9	2	US-08-287-537B-3	Sequence 3, Appli
41	55	68.8	9	2	US-08-483-898-19	Sequence 19, Appl
42	55	68.8	9	3	US-09-087-716-19	Sequence 19, Appl
43	55	68.8	9	3	US-08-853-733B-28	Sequence 28, Appl
44	55	68.8	9	3	US-08-554-385-6	Sequence 6, Appli
45	55	68.8	9	3	US-09-157-753-19	Sequence 19, Appl

ALIGNMENTS

RESULT 1
US-09-208-966-15
; Sequence 15, Application US/09208966
; Patent No. 6221355
; GENERAL INFORMATION:
; APPLICANT: Dowdy, Steven F.
; TITLE OF INVENTION: ANTI-PATHOGEN SYSTEM AND METHODS OF USE THEREOF
; FILE REFERENCE: 48881/1742
; CURRENT APPLICATION NUMBER: US/09/208,966
; CURRENT FILING DATE: 1998-12-10
; EARLIER APPLICATION NUMBER: 60/082,402
; EARLIER FILING DATE: 1998-04-20
; EARLIER APPLICATION NUMBER: 60/069,012
; EARLIER FILING DATE: 1997-12-10
; NUMBER OF SEQ ID NOS: 57
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 15
; LENGTH: 16
; TYPE: PRT
; ORGANISM: human
US-09-208-966-15

Query Match 81.2%; Score 65; DB 4; Length 16;
Best Local Similarity 100.0%; Pred. No. 0.00034;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 YPYDVPDYAGS 11
||||||||||
Db 4 YPYDVPDYAGS 14

RESULT 2
US-08-789-333F-51
; Sequence 51, Application US/08789333F
; Patent No. 6153380
; GENERAL INFORMATION:
; APPLICANT: No. 6153380an, Garry P
; APPLICANT: Rothenberg, S. M.
; TITLE OF INVENTION: METHODS FOR SCREENING FOR TRANSDOMINANT INTRACELLULAR
; TITLE OF INVENTION: EFFECTOR PEPTIDES AND RNA MOLECULES
; FILE REFERENCE: A642601DJBRMSDSS
; CURRENT APPLICATION NUMBER: US/08/789,333F
; CURRENT FILING DATE: 1997-01-23
; PRIOR APPLICATION NUMBER: 08/589,108
; PRIOR FILING DATE: 1996-01-23
; PRIOR APPLICATION NUMBER: 08/589,911
; PRIOR FILING DATE: 1996-01-23
; NUMBER OF SEQ ID NOS: 102
; SOFTWARE: PatentIn Ver. 2.0

```
; SEQ ID NO 51
;   LENGTH: 17
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: epitope tag
;   OTHER INFORMATION: sequence.
US-08-789-333F-51


  Query Match             81.2%;  Score 65;  DB 4;  Length 17;
  Best Local Similarity   100.0%;  Pred. No. 0.00037;
  Matches   11;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 YPYDVPDYAGS 11
          ||||||||||||
Db      5 YPYDVPDYAGS 15


RESULT   3
US-08-634-060-27
; Sequence 27, Application US/08634060
; Patent No. 5712136
;  GENERAL INFORMATION:
;    APPLICANT:  Wickham, Thomas J.
;    APPLICANT:  Kovesdi, Imre
;    APPLICANT:  Roelvink, Petrus W.
;    TITLE OF INVENTION:  ADENOVIRAL-MEDIATED CELL TARGETING COMMANDED BY
;    TITLE OF INVENTION:  THE ADENOVIRUS PENTON BASE PROTEIN
;    NUMBER OF SEQUENCES:  60
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Leydig, Voit & Mayer, Ltd.
;      STREET:  Two Prudential Plaza, Suite 4900
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  USA
;      ZIP:  60601
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/634,060
;      FILING DATE:
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/303,162
;      FILING DATE:  08-SEP-1994
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Kilyk, John Jr.
;      REGISTRATION NUMBER:  30763
;      REFERENCE/DOCKET NUMBER:  71602
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (312) 616-5600
;      TELEFAX:  (312) 616-5700
;  INFORMATION FOR SEQ ID NO:  27:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  25 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-634-060-27


  Query Match             81.2%;  Score 65;  DB 1;  Length 25;
  Best Local Similarity   100.0%;  Pred. No. 0.00055;
  Matches   11;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 YPYDVPDYAGS 11
          ||||||||||||
Db     14 YPYDVPDYAGS 24


RESULT   4
US-08-700-846-10
; Sequence 10, Application US/08700846
; Patent No. 5962311
;  GENERAL INFORMATION:
;    APPLICANT:  WICKHAM, THOMAS J.
;    APPLICANT:  ROELVINK, PETRUS W.
;    APPLICANT:  KOVESDI, IMRE
;    TITLE OF INVENTION:  A SHORT-SHAFTED ADENOVIRAL FIBER AND ITS
;    TITLE OF INVENTION:  USE
;    NUMBER OF SEQUENCES:  17
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  LEYDIG, VOIT & MAYER, LTD.
;      STREET:  TWO PRUDENTIAL PLAZA, SUITE 4900
;      CITY:  CHICAGO
;      STATE:  IL
;      COUNTRY:  USA
;      ZIP:  60601-6780
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/700,846
;      FILING DATE:  21-AUG-1996
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  LARCHER, CAROL
;      REGISTRATION NUMBER:  35243
;      REFERENCE/DOCKET NUMBER:  74294
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (312) 616-5600
;      TELEFAX:  (312) 616-5700
;  INFORMATION FOR SEQ ID NO:  10:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  25 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-700-846-10


  Query Match             81.2%;  Score 65;  DB 2;  Length 25;
  Best Local Similarity   100.0%;  Pred. No. 0.00055;
  Matches   11;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 YPYDVPDYAGS 11
          ||||||||||||
Db     14 YPYDVPDYAGS 24


RESULT   5
US-08-321-071A-5
; Sequence 5, Application US/08321071A
; Patent No. 5672686
;  GENERAL INFORMATION:
;    APPLICANT:  CHITTENDEN, Thomas D.
;    TITLE OF INVENTION:  APOPTOSIS RELATED PROTEIN Bcl-Y, AND METHODS
;    TITLE OF INVENTION:  OF USE THEREOF
;    NUMBER OF SEQUENCES:  31
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Hale and Dorr
;      STREET:  1455 Pennsylvania Avenue, N.W.
;      CITY:  Washington
;      STATE:  D.C.
;      ZIP:  20004
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
```

```
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/321,071A
;      FILING DATE:  11-OCT-1994
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US95/10103
;      FILING DATE:  09-AUG-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/287,427
;      FILING DATE:  09-AUG-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  WIXON, HENRY N.
;      REGISTRATION NUMBER:  32,073
;      REFERENCE/DOCKET NUMBER:  104322.121CIP
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  202-942-8400
;      TELEFAX:  202-942-8484
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  15 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-321-071A-5


  Query Match             78.8%;  Score 63;  DB 1;  Length 15;
  Best Local Similarity   84.6%;  Pred. No. 0.00066;
  Matches   11;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 YPYDVPDYAGSGS 13
           |||||||||  ||
Db       3 YPYDVPDYASLGS 15


RESULT   6
US-08-902-516-2
; Sequence 2, Application US/08902516
; Patent No. 5891432
;  GENERAL INFORMATION:
;    APPLICANT:  Soo Hoo, William
;    TITLE OF INVENTION:  MEMBRANE-BOUND CYTOKINE COMPOSITIONS
;    TITLE OF INVENTION:  COMPRISING GM-CSF AND METHODS OF MODULATING AN IMMUNE
;    TITLE OF INVENTION:  RESPONSE USING SAME
;    NUMBER OF SEQUENCES:  50
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  CAMPBELL & FLORES, LLP
;      STREET:  4370 La Jolla Village Drive, Suite 700
;      CITY:  San Diego
;      STATE:  California
;      COUNTRY:  United States
;      ZIP:  92121
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/902,516
;      FILING DATE:  29-JUL-1997
;      CLASSIFICATION:  424
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Campbell, Cathryn A.
;      REGISTRATION NUMBER:  31,815
;      REFERENCE/DOCKET NUMBER:  P-IM 2442
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (619)535-9001
;      TELEFAX:  (619)535-8949
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  219 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-902-516-2


  Query Match             77.5%;  Score 62;  DB 2;  Length 219;
  Best Local Similarity   90.9%;  Pred. No. 0.016;
  Matches   10;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 YPYDVPDYAGS 11
           ||||||||||:
Db      22 YPYDVPDYAGA 32


RESULT   7
US-08-990-140-13
; Sequence 13, Application US/08990140A
; Patent No. 6093795
; GENERAL INFORMATION:
;  APPLICANT: Olsen, Henrik S.
;  APPLICANT: Ruben, Steven M.
;  APPLICANT: Sonenberg, Nahum
;  APPLICANT: Methot, Nathalie
;  APPLICANT: Rom, Eran
;  TITLE OF INVENTION: Human Prt1-like Subunit Protein (hPrt1) and Human
;  TITLE OF INVENTION: eIF4GF-like Protein (p97) Genes
;  FILE REFERENCE: 1488.0700001
;  CURRENT APPLICATION NUMBER: US/08/990,140A
;  CURRENT FILING DATE: 1997-12-12
;  EARLIER APPLICATION NUMBER: US 60/033,151
;  EARLIER FILING DATE: 1996-12-13
;  NUMBER OF SEQ ID NOS: 13
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 13
;   LENGTH: 10
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-08-990-140-13


  Query Match             76.2%;  Score 61;  DB 3;  Length 10;
  Best Local Similarity   100.0%;  Pred. No. 0.00088;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 YPYDVPDYAG 10
           ||||||||||
Db       1 YPYDVPDYAG 10


RESULT   8
US-08-282-951-2
; Sequence 2, Application US/08282951
; Patent No. 5665590
;  GENERAL INFORMATION:
;    APPLICANT:  YANG, ZHI
;    TITLE OF INVENTION:  METHOD FOR ISOLATING AND DIRECTLY
;    TITLE OF INVENTION:  CLONING GENES WHICH ENCODE CELL-SURFACE AND SECRETED
;    TITLE OF INVENTION:  PROTEINS
;    NUMBER OF SEQUENCES:  12
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Morrison & Foerster
;      STREET:  755 Page Mill Road
;      CITY:  Palo Alto
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94304-1018
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
```

```
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/282,951
;      FILING DATE:  29-JUL-1994
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  LEHNHARDT, SUSAN K.
;      REGISTRATION NUMBER:  33,943
;      REFERENCE/DOCKET NUMBER:  20296-20012.00
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 813-5600
;      TELEFAX:  (415) 494-0792
;      TELEX:  706141
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  102 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-282-951-2


  Query Match             75.6%;  Score 60.5;  DB 1;  Length 102;
  Best Local Similarity   75.0%;  Pred. No. 0.012;
  Matches   12;  Conservative    1;  Mismatches    0;  Indels    3;  Gaps    1;

Qy       1 YPYDVPDYA---GSGS 13
           |||||||||   |||:
Db      56 YPYDVPDYAPNKGSGT 71


RESULT  9
US-08-282-951-6
; Sequence 6, Application US/08282951
; Patent No. 5665590
;  GENERAL INFORMATION:
;    APPLICANT:  YANG, ZHI
;    TITLE OF INVENTION:  METHOD FOR ISOLATING AND DIRECTLY
;    TITLE OF INVENTION:  CLONING GENES WHICH ENCODE CELL-SURFACE AND SECRETED
;    TITLE OF INVENTION:  PROTEINS
;    NUMBER OF SEQUENCES:  12
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Morrison & Foerster
;      STREET:  755 Page Mill Road
;      CITY:  Palo Alto
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94304-1018
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/282,951
;      FILING DATE:  29-JUL-1994
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  LEHNHARDT, SUSAN K.
;      REGISTRATION NUMBER:  33,943
;      REFERENCE/DOCKET NUMBER:  20296-20012.00
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 813-5600
;      TELEFAX:  (415) 494-0792
;      TELEX:  706141
;  INFORMATION FOR SEQ ID NO:  6:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  272 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-282-951-6


  Query Match             75.6%;  Score 60.5;  DB 1;  Length 272;
  Best Local Similarity   75.0%;  Pred. No. 0.035;
  Matches   12;  Conservative    1;  Mismatches    0;  Indels    3;  Gaps    1;

Qy       1 YPYDVPDYA---GSGS 13
           |||||||||   |||:
Db     227 YPYDVPDYAPNKGSGT 242


RESULT  10
US-08-723-726-1
; Sequence 1, Application US/08723726
; Patent No. 5851521
;  GENERAL INFORMATION:
;    APPLICANT:  BRANELLEC, Didier
;    APPLICANT:  WALSH, Kenneth
;    APPLICANT:  ISNER, Jeffrey M.
;    APPLICANT:  DENEFLE, Patrice
;    TITLE OF INVENTION:  VIRAL VECTORS AND THEIR USE FOR TREATING
;    TITLE OF INVENTION:  HYPERPROLIFERATIVE DISORDERS, IN PARTICULAR RESTENOSIS
;    NUMBER OF SEQUENCES:  3
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Rhone-Poulenc Rorer Inc.
;      STREET:  Mail Drop 3C43, P.O. Box 5093
;      CITY:  Collegeville
;      STATE:  PA
;      COUNTRY:  USA
;      ZIP:  19426-0997
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/723,726
;      FILING DATE:  30-SEP-1996
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  WO PCT/US96/04493
;      FILING DATE:  28-MAR-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  FR 95-04234
;      FILING DATE:  31-MAR-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Savitzky, Martin F.
;      REGISTRATION NUMBER:  29,699
;      REFERENCE/DOCKET NUMBER:  ST95022A-US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (610)454-3816
;      TELEFAX:  (610)454-3808
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  14 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    FRAGMENT TYPE:  internal
US-08-723-726-1


  Query Match             73.8%;  Score 59;  DB 2;  Length 14;
  Best Local Similarity   83.3%;  Pred. No. 0.0026;
  Matches   10;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 YPYDVPDYAGSG 12
           |||||||||  |
Db       1 YPYDVPDYASLG 12


RESULT  11
US-08-390-874C-13
```

```
; Sequence 13, Application US/08390874C
; Patent No. 6043062
;  GENERAL INFORMATION:
;    APPLICANT:  Klippel, Anke
;    APPLICANT:  Williams, Lewis T.
;    TITLE OF INVENTION:  A Constitutively Active
;    TITLE OF INVENTION:  Phosphatidylinositol 3-Kinase and Uses Thereof
;    NUMBER OF SEQUENCES:  13
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Townsend and Townsend and Crew LLP
;      STREET:  Two Embarcadero Center, Eighth Floor
;      CITY:  San Francisco
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94111-3834
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/390,874C
;      FILING DATE:  17-FEB-1995
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Dow, Karen B.
;      REGISTRATION NUMBER:  29,684
;      REFERENCE/DOCKET NUMBER:  02307K-057000US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 576-0200
;      TELEFAX:  (415) 576-0300
;  INFORMATION FOR SEQ ID NO:  13:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  16 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-390-874C-13


  Query Match             73.8%;  Score 59;  DB 3;  Length 16;
  Best Local Similarity   83.3%;  Pred. No. 0.0029;
  Matches   10;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 YPYDVPDYAGSG 12
           |||||||||  |
Db       2 YPYDVPDYASLG 13


RESULT  12
US-09-265-772-13
; Sequence 13, Application US/09265772
; Patent No. 6300111
;  GENERAL INFORMATION:
;    APPLICANT:  Klippel, Anke
;    APPLICANT:  Williams, Lewis T.
;    TITLE OF INVENTION:  A Constitutively Active
;    TITLE OF INVENTION:  Phosphatidylinositol 3-Kinase and Uses Thereof
;    NUMBER OF SEQUENCES:  13
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Townsend and Townsend and Crew LLP
;      STREET:  Two Embarcadero Center, Eighth Floor
;      CITY:  San Francisco
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94111-3834
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/265,772
;      FILING DATE:  10-MAR-1999
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/390,874
;      FILING DATE:  17-FEB-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Dow, Karen B.
;      REGISTRATION NUMBER:  29,684
;      REFERENCE/DOCKET NUMBER:  02307K-057010US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 576-0200
;      TELEFAX:  (415) 576-0300
;  INFORMATION FOR SEQ ID NO:  13:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  16 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-09-265-772-13


  Query Match             73.8%;  Score 59;  DB 4;  Length 16;
  Best Local Similarity   83.3%;  Pred. No. 0.0029;
  Matches   10;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 YPYDVPDYAGSG 12
           |||||||||  |
Db       2 YPYDVPDYASLG 13


RESULT  13
US-09-169-015-52
; Sequence 52, Application US/09169015
; Patent No. 6180343
; GENERAL INFORMATION:
;  APPLICANT:  Anderson, David
;  APPLICANT:  Bogenberger, Jakob M.
;  TITLE OF INVENTION:  Green Fluorescent Protein Fusions with Random Peptides
;  FILE REFERENCE:  A66900/DJB/RMS/SJR
;  CURRENT APPLICATION NUMBER: US/09/169,015
;  CURRENT FILING DATE:  1998-10-08
;  NUMBER OF SEQ ID NOS:  55
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 52
;   LENGTH: 19
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:  synthetic
US-09-169-015-52


  Query Match             73.8%;  Score 59;  DB 4;  Length 19;
  Best Local Similarity   83.3%;  Pred. No. 0.0035;
  Matches   10;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 YPYDVPDYAGSG 12
           |||||||||  |
Db       5 YPYDVPDYASLG 16


RESULT  14
US-08-672-213-39
; Sequence 39, Application US/08672213
; Patent No. 6306649
;  GENERAL INFORMATION:
;    APPLICANT:  GILMAN, Michael Z.
;    APPLICANT:  NATESAN, Sridaran
;    TITLE OF INVENTION:  USE OF HETEROLOGOUS TRANSCRIPTION
;    TITLE OF INVENTION:  FACTORS IN GENE THERAPY
;    NUMBER OF SEQUENCES:  72
```

```
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  ARIAD Gene Therapeutics, Inc.
;      STREET:  26 Landsdowne Street
;      CITY:  Cambridge
;      STATE:  Massachusetts
;      COUNTRY:  USA
;      ZIP:  02139-4234
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/672,213
;      FILING DATE:  27-JUN-1996
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/000,553
;      FILING DATE:  27-JUN-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/019,614
;      FILING DATE:  29-DEC-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  BERSTEIN, David L.
;      REGISTRATION NUMBER:  31,235
;      REFERENCE/DOCKET NUMBER:  ARIAD 346B
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617-494-0400
;      TELEFAX:  617-494-0208
;  INFORMATION FOR SEQ ID NO:  39:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  27 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-672-213-39


  Query Match           73.8%;  Score 59;  DB 4;  Length 27;
  Best Local Similarity  83.3%;  Pred. No. 0.0051;
  Matches   10;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 YPYDVPDYAGSG 12
           |||||||||  |
Db       5 YPYDVPDYASLG 16


RESULT  15
US-08-672-213-29
; Sequence 29, Application US/08672213
; Patent No. 6306649
;  GENERAL INFORMATION:
;    APPLICANT:  GILMAN, Michael Z.
;    APPLICANT:  NATESAN, Sridaran
;    TITLE OF INVENTION:  USE OF HETEROLOGOUS TRANSCRIPTION
;    TITLE OF INVENTION:  FACTORS IN GENE THERAPY
;    NUMBER OF SEQUENCES:  72
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  ARIAD Gene Therapeutics, Inc.
;      STREET:  26 Landsdowne Street
;      CITY:  Cambridge
;      STATE:  Massachusetts
;      COUNTRY:  USA
;      ZIP:  02139-4234
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/672,213
;      FILING DATE:  27-JUN-1996
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/000,553
;      FILING DATE:  27-JUN-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/019,614
;      FILING DATE:  29-DEC-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  BERSTEIN, David L.
;      REGISTRATION NUMBER:  31,235
;      REFERENCE/DOCKET NUMBER:  ARIAD 346B
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617-494-0400
;      TELEFAX:  617-494-0208
;  INFORMATION FOR SEQ ID NO:  29:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  40 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-672-213-29


  Query Match           73.8%;  Score 59;  DB 4;  Length 40;
  Best Local Similarity  83.3%;  Pred. No. 0.0078;
  Matches   10;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 YPYDVPDYAGSG 12
           |||||||||  |
Db       5 YPYDVPDYASLG 16


Search completed: April 11, 2002, 10:14:44
Job time: 34 sec
```

GenCore version 4.5

OM protein - protein search, using sw model

Run on: April 11, 2002, 10:16:16 ; Search time 27.06 Seconds
(without alignments)
39.410 Million cell updates/sec

Title: US-09-284-787-2
Perfect score: 80
Sequence: 1 YPYDVPDYAGSGSK 14

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_68:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	55	68.8	330	2	JQ2374	hemagglutinin - in
2	55	68.8	330	2	JQ2375	hemagglutinin - in
3	55	68.8	331	2	JQ2377	hemagglutinin - in
4	55	68.8	331	2	JQ2378	hemagglutinin - in
5	55	68.8	347	2	S52173	hemagglutinin - in
6	55	68.8	347	2	S52174	hemagglutinin - in
7	55	68.8	347	2	S52175	hemagglutinin - in
8	55	68.8	347	2	S52176	hemagglutinin - in
9	55	68.8	347	2	S52178	hemagglutinin - in
10	55	68.8	347	2	S52179	hemagglutinin - in
11	55	68.8	347	2	S52180	hemagglutinin - in
12	55	68.8	347	2	S52182	hemagglutinin - in
13	55	68.8	347	2	S52183	hemagglutinin - in
14	55	68.8	347	2	S52185	hemagglutinin - in
15	55	68.8	347	2	S52186	hemagglutinin - in
16	55	68.8	347	2	S52187	hemagglutinin - in
17	55	68.8	347	2	S52188	hemagglutinin - in
18	55	68.8	347	2	S52189	hemagglutinin - in
19	55	68.8	347	2	S52191	hemagglutinin - in
20	55	68.8	347	2	S52195	hemagglutinin - in
21	55	68.8	347	2	S52197	hemagglutinin - in
22	55	68.8	347	2	S52199	hemagglutinin - in
23	55	68.8	362	2	S38637	hemagglutinin - in
24	55	68.8	550	1	HMIVS2	hemagglutinin prec
25	55	68.8	550	1	HMIVS3	hemagglutinin prec
26	55	68.8	550	1	HMIV77	hemagglutinin prec
27	55	68.8	550	1	HMIV80	hemagglutinin prec
28	55	68.8	550	1	HMIV33	hemagglutinin prec
29	55	68.8	550	1	HMIV89	hemagglutinin prec
30	55	68.8	550	1	HMIV98	hemagglutinin prec
31	55	68.8	550	1	HMIV15	hemagglutinin prec
32	55	68.8	550	1	HMIV86	hemagglutinin prec
33	55	68.8	550	2	JQ1153	hemagglutinin prec
34	55	68.8	550	2	JQ1154	hemagglutinin prec
35	55	68.8	550	2	JQ1155	hemagglutinin prec
36	55	68.8	550	2	JQ1156	hemagglutinin prec
37	55	68.8	565	1	HMIVE2	hemagglutinin prec
38	55	68.8	565	1	HMIVE4	hemagglutinin prec
39	55	68.8	566	1	HMIVH	hemagglutinin prec
40	55	68.8	566	1	HMIVHA	hemagglutinin prec
41	55	68.8	566	1	HMIVHM	hemagglutinin prec
42	55	68.8	566	1	HMIV6	hemagglutinin prec
43	55	68.8	567	1	HMIVV	hemagglutinin prec
44	54	67.5	565	1	HMIVE1	hemagglutinin prec
45	54	67.5	565	1	HMIVE5	hemagglutinin prec

ALIGNMENTS

RESULT 1
JQ2374
hemagglutinin - influenza A virus (strain SN1289)
C;Species: influenza A virus
C;Date: 14-Jul-1994 #sequence_revision 14-Jul-1994 #text_change 07-May-1999
C;Accession: JQ2374
R;Rocha, E.P.; Xu, X.; Hall, H.E.; Allen, J.R.; Regnery, H.L.; Cox, N.J.
J. Gen. Virol. 74, 2513-2518, 1993
A;Title: Comparison of 10 influenza A (H1N1 and H3N2) haemagglutinin sequences obtain
A;Reference number: JQ2369; MUID:94065682
A;Accession: JQ2374
A;Molecule type: mRNA
A;Residues: 1-330 <ROC>
A;Cross-references: GB:L20115
A;Experimental source: subtype H3N2
C;Superfamily: influenza virus hemagglutinin

Query Match 68.8%; Score 55; DB 2; Length 330;
Best Local Similarity 100.0%; Pred. No. 0.14;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy      1 YPYDVPDYA 9
          |||||||||
Db     98 YPYDVPDYA 106
```

RESULT 2
JQ2375
hemagglutinin - influenza A virus (strain SN1389 and IN0190)
C;Species: influenza A virus
C;Date: 14-Jul-1994 #sequence_revision 14-Jul-1994 #text_change 07-May-1999
C;Accession: JQ2375; JQ2376
R;Rocha, E.P.; Xu, X.; Hall, H.E.; Allen, J.R.; Regnery, H.L.; Cox, N.J.
J. Gen. Virol. 74, 2513-2518, 1993
A;Title: Comparison of 10 influenza A (H1N1 and H3N2) haemagglutinin sequences obtain
A;Reference number: JQ2369; MUID:94065682
A;Accession: JQ2375
A;Molecule type: mRNA
A;Residues: 1-330 <ROC>
A;Cross-references: GB:L20119
A;Experimental source: subtype H3N2
C;Superfamily: influenza virus hemagglutinin

Query Match 68.8%; Score 55; DB 2; Length 330;
Best Local Similarity 100.0%; Pred. No. 0.14;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy      1 YPYDVPDYA 9
          |||||||||
Db     98 YPYDVPDYA 106
```

RESULT 3
JQ2377
hemagglutinin - influenza A virus (strain SD0191)
C;Species: influenza A virus
C;Date: 14-Jul-1994 #sequence_revision 14-Jul-1994 #text_change 07-May-1999
C;Accession: JQ2377
R;Rocha, E.P.; Xu, X.; Hall, H.E.; Allen, J.R.; Regnery, H.L.; Cox, N.J.
J. Gen. Virol. 74, 2513-2518, 1993
A;Title: Comparison of 10 influenza A (H1N1 and H3N2) haemagglutinin sequences obtained
A;Reference number: JQ2369; MUID:94065682
A;Accession: JQ2377
A;Molecule type: mRNA
A;Residues: 1-331 <ROC>
A;Cross-references: GB:L20105
A;Experimental source: subtype H3N2
C;Superfamily: influenza virus hemagglutinin

Query Match 68.8%; Score 55; DB 2; Length 331;
Best Local Similarity 100.0%; Pred. No. 0.15;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 YPYDVPDYA 9
|||||||||
Db 98 YPYDVPDYA 106

RESULT 4
JQ2378
hemagglutinin - influenza A virus (strain HI0191)
C;Species: influenza A virus
C;Date: 14-Jul-1994 #sequence_revision 14-Jul-1994 #text_change 07-May-1999
C;Accession: JQ2378
R;Rocha, E.P.; Xu, X.; Hall, H.E.; Allen, J.R.; Regnery, H.L.; Cox, N.J.
J. Gen. Virol. 74, 2513-2518, 1993
A;Title: Comparison of 10 influenza A (H1N1 and H3N2) haemagglutinin sequences obtained
A;Reference number: JQ2369; MUID:94065682
A;Accession: JQ2378
A;Molecule type: mRNA
A;Residues: 1-331 <ROC>
A;Cross-references: GB:L20102
A;Experimental source: subtype H3N2
C;Superfamily: influenza virus hemagglutinin

Query Match 68.8%; Score 55; DB 2; Length 331;
Best Local Similarity 100.0%; Pred. No. 0.15;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 YPYDVPDYA 9
|||||||||
Db 98 YPYDVPDYA 106

RESULT 5
S52173
hemagglutinin - influenza A virus (fragment)
C;Species: influenza A virus
C;Date: 07-May-1995 #sequence_revision 23-Feb-1996 #text_change 23-Mar-2001
C;Accession: S52173; S52190; S52192; S52177; S52181; S52194; S52196; S52184; S52193
R;Ellis, J.S.
submitted to the EMBL Data Library, October 1994
A;Description: Antigenic and genetic variation in the haemagglutinin of recently circula
A;Reference number: S52173
A;Accession: S52173
A;Molecule type: mRNA
A;Residues: 1-347 <ELL>
A;Cross-references: EMBL:Z46392; NID:g609020; PIDN:CAA86527.1; PID:g609021
A;Experimental source: isolate A/Beijing/32/92
A;Accession: S52190
A;Molecule type: mRNA
A;Residues: 1-347 <EL2>
A;Cross-references: EMBL:Z46410; NID:g609053; PIDN:CAA86545.1; PID:g609054
A;Experimental source: isolate A/Hong Kong/23/92
A;Accession: S52192
A;Molecule type: mRNA
A;Residues: 1-347 <EL3>
A;Cross-references: EMBL:Z46409; NID:g609057; PIDN:CAA86544.1; PID:g609058
A;Experimental source: isolate A/Hong Kong/34/90
A;Accession: S52177
A;Molecule type: mRNA
A;Residues: 1-134,'K',136-156,'L',158-173,'F',175-188,'S',190-218,'F',220-225,'L',227
A;Cross-references: EMBL:Z46395; NID:g609027; PIDN:CAA86530.1; PID:g1228087
A;Experimental source: isolate A/England/269/93
A;Accession: S52181
A;Molecule type: mRNA
A;Residues: 1-134,'K',136-156,'L',158-173,'F',175-188,'S',190-218,'F',220-225,'L',227
A;Cross-references: EMBL:Z46399; NID:g609035; PIDN:CAA86534.1; PID:g609036
A;Experimental source: isolate A/England/328/93
A;Accession: S52194
A;Molecule type: mRNA
A;Residues: 1-134,'K',136-156,'L',158-173,'F',175-188,'S',190-218,'F',220-225,'L',227
A;Cross-references: EMBL:Z46414; NID:g609061; PIDN:CAA86549.1; PID:g609062
A;Experimental source: isolate A/Scotland/160/93
A;Accession: S52196
A;Molecule type: mRNA
A;Residues: 1-134,'K',136-156,'L',158-173,'F',175-188,'S',190-218,'F',220-225,'L',227
A;Cross-references: EMBL:Z46415; NID:g609065; PIDN:CAA86550.1; PID:g609066
A;Experimental source: isolate A/Scotland/174/93
A;Accession: S52184
A;Molecule type: mRNA
A;Residues: 1-74,'N',76-120,'T',122-156,'L',158-173,'F',175-188,'S',190-200,'K',202-2
A;Cross-references: EMBL:Z46402; NID:g609041; PIDN:CAA86537.1; PID:g609042
A;Experimental source: isolate A/England/471/93
A;Accession: S52193
A;Molecule type: mRNA
A;Residues: 1-74,'N',76-120,'T',122-156,'L',158-173,'F',175-188,'S',190-200,'K',202-2
A;Cross-references: EMBL:Z46413; NID:g609059; PIDN:CAA86548.1; PID:g609060
A;Experimental source: isolate A/Scotland/142/93
C;Superfamily: influenza virus hemagglutinin

Query Match 68.8%; Score 55; DB 2; Length 347;
Best Local Similarity 100.0%; Pred. No. 0.15;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 YPYDVPDYA 9
|||||||||
Db 98 YPYDVPDYA 106

RESULT 6
S52174
hemagglutinin - influenza A virus (isolate A/Beijing/353/89) (fragment)
C;Species: influenza A virus
A;Variety: isolate A/Beijing/353/89
C;Date: 07-May-1995 #sequence_revision 23-Feb-1996 #text_change 20-Sep-1999
C;Accession: S52174
R;Ellis, J.S.
submitted to the EMBL Data Library, October 1994
A;Description: Antigenic and genetic variation in the haemagglutinin of recently circ
A;Reference number: S52173
A;Accession: S52174
A;Molecule type: mRNA
A;Residues: 1-347 <ELL>
A;Cross-references: EMBL:Z46391; NID:g609022; PIDN:CAA86526.1; PID:g609023
A;Experimental source: isolate A/Beijing/353/89
C;Superfamily: influenza virus hemagglutinin

Query Match 68.8%; Score 55; DB 2; Length 347;
Best Local Similarity 100.0%; Pred. No. 0.15;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       1 YPYDVPDYA 9
           |||||||||
Db      98 YPYDVPDYA 106


RESULT   7
S52175
hemagglutinin - influenza A virus (isolate A/England/1/93) (fragment)
C;Species: influenza A virus
A;Variety: isolate A/England/1/93
C;Date: 07-May-1995 #sequence_revision 23-Feb-1996 #text_change 23-Mar-2001
C;Accession: S52175
R;Ellis, J.S.
submitted to the EMBL Data Library, October 1994
A;Description: Antigenic and genetic variation in the haemagglutinin of recently circula
A;Reference number: S52173
A;Accession: S52175
A;Molecule type: mRNA
A;Residues: 1-347 <ELL>
A;Cross-references: EMBL:Z46393; NID:g609024; PIDN:CAA86528.1; PID:g940547
A;Experimental source: isolate A/England/1/93
C;Superfamily: influenza virus hemagglutinin


  Query Match             68.8%;  Score 55;  DB 2;  Length 347;
  Best Local Similarity   100.0%;  Pred. No. 0.15;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 YPYDVPDYA 9
           |||||||||
Db      98 YPYDVPDYA 106


RESULT   8
S52176
hemagglutinin - influenza A virus (isolate A/England/247/93) (fragment)
C;Species: influenza A virus
A;Variety: isolate A/England/247/93
C;Date: 07-May-1995 #sequence_revision 23-Feb-1996 #text_change 23-Mar-2001
C;Accession: S52176
R;Ellis, J.S.
submitted to the EMBL Data Library, October 1994
A;Description: Antigenic and genetic variation in the haemagglutinin of recently circula
A;Reference number: S52173
A;Accession: S52176
A;Molecule type: mRNA
A;Residues: 1-347 <ELL>
A;Cross-references: EMBL:Z46394; NID:g609025; PIDN:CAA86529.1; PID:g609026
A;Experimental source: isolate A/England/247/93
C;Superfamily: influenza virus hemagglutinin


  Query Match             68.8%;  Score 55;  DB 2;  Length 347;
  Best Local Similarity   100.0%;  Pred. No. 0.15;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 YPYDVPDYA 9
           |||||||||
Db      98 YPYDVPDYA 106


RESULT   9
S52178
hemagglutinin - influenza A virus (isolate A/England/284/93) (fragment)
C;Species: influenza A virus
A;Variety: isolate A/England/284/93
C;Date: 07-May-1995 #sequence_revision 23-Feb-1996 #text_change 23-Mar-2001
C;Accession: S52178
R;Ellis, J.S.
submitted to the EMBL Data Library, October 1994
A;Description: Antigenic and genetic variation in the haemagglutinin of recently circula
A;Reference number: S52173
A;Accession: S52178
A;Molecule type: mRNA
A;Residues: 1-347 <ELL>
A;Cross-references: EMBL:Z46396; NID:g609029; PIDN:CAA86531.1; PID:g609030
A;Experimental source: isolate A/England/284/93
C;Superfamily: influenza virus hemagglutinin


  Query Match             68.8%;  Score 55;  DB 2;  Length 347;
  Best Local Similarity   100.0%;  Pred. No. 0.15;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 YPYDVPDYA 9
           |||||||||
Db      98 YPYDVPDYA 106


RESULT  10
S52179
hemagglutinin - influenza A virus (isolate A/England/286/93) (fragment)
C;Species: influenza A virus
A;Variety: isolate A/England/286/93
C;Date: 07-May-1995 #sequence_revision 23-Feb-1996 #text_change 23-Mar-2001
C;Accession: S52179
R;Ellis, J.S.
submitted to the EMBL Data Library, October 1994
A;Description: Antigenic and genetic variation in the haemagglutinin of recently circ
A;Reference number: S52173
A;Accession: S52179
A;Molecule type: mRNA
A;Residues: 1-347 <ELL>
A;Cross-references: EMBL:Z46397; NID:g609031; PIDN:CAA86532.1; PID:g609032
A;Experimental source: isolate A/England/286/93
C;Superfamily: influenza virus hemagglutinin


  Query Match             68.8%;  Score 55;  DB 2;  Length 347;
  Best Local Similarity   100.0%;  Pred. No. 0.15;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 YPYDVPDYA 9
           |||||||||
Db      98 YPYDVPDYA 106


RESULT  11
S52180
hemagglutinin - influenza A virus (isolate A/England/289/93) (fragment)
C;Species: influenza A virus
A;Variety: isolate A/England/289/93
C;Date: 07-May-1995 #sequence_revision 23-Feb-1996 #text_change 23-Mar-2001
C;Accession: S52180
R;Ellis, J.S.
submitted to the EMBL Data Library, October 1994
A;Description: Antigenic and genetic variation in the haemagglutinin of recently circ
A;Reference number: S52173
A;Accession: S52180
A;Molecule type: mRNA
A;Residues: 1-347 <ELL>
A;Cross-references: EMBL:Z46398; NID:g609033; PIDN:CAA86533.1; PID:g609034
A;Experimental source: isolate A/England/289/93
C;Superfamily: influenza virus hemagglutinin


  Query Match             68.8%;  Score 55;  DB 2;  Length 347;
  Best Local Similarity   100.0%;  Pred. No. 0.15;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 YPYDVPDYA 9
           |||||||||
Db      98 YPYDVPDYA 106


RESULT  12
```

S52182
hemagglutinin - influenza A virus (isolate A/England/328/93) (fragment)
C;Species: influenza A virus
A;Variety: isolate A/England/328/93
C;Date: 07-May-1995 #sequence_revision 23-Feb-1996 #text_change 23-Mar-2001
C;Accession: S52182
R;Ellis, J.S.
submitted to the EMBL Data Library, October 1994
A;Description: Antigenic and genetic variation in the haemagglutinin of recently circula
A;Reference number: S52173
A;Accession: S52182
A;Molecule type: mRNA
A;Residues: 1-347 <ELL>
A;Cross-references: EMBL:Z46400; NID:g609037; PIDN:CAA86535.1; PID:g609038
A;Experimental source: isolate A/England/346/93
C;Superfamily: influenza virus hemagglutinin

Query Match 68.8%; Score 55; DB 2; Length 347;
Best Local Similarity 100.0%; Pred. No. 0.15;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 YPYDVPDYA 9
||||||||| |
Db 98 YPYDVPDYA 106

RESULT 13
S52183
hemagglutinin - influenza A virus (isolate A/England/347/93) (fragment)
C;Species: influenza A virus
A;Variety: isolate A/England/347/93
C;Date: 07-May-1995 #sequence_revision 23-Feb-1996 #text_change 23-Mar-2001
C;Accession: S52183
R;Ellis, J.S.
submitted to the EMBL Data Library, October 1994
A;Description: Antigenic and genetic variation in the haemagglutinin of recently circula
A;Reference number: S52173
A;Accession: S52183
A;Molecule type: mRNA
A;Residues: 1-347 <ELL>
A;Cross-references: EMBL:Z46401; NID:g609039; PIDN:CAA86536.1; PID:g609040
A;Experimental source: isolate A/England/347/93
C;Superfamily: influenza virus hemagglutinin

Query Match 68.8%; Score 55; DB 2; Length 347;
Best Local Similarity 100.0%; Pred. No. 0.15;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 YPYDVPDYA 9
||||||||| |
Db 98 YPYDVPDYA 106

RESULT 14
S52185
hemagglutinin - influenza A virus (isolate A/England/67/94) (fragment)
C;Species: influenza A virus
A;Variety: isolate A/England/67/94
C;Date: 07-May-1995 #sequence_revision 23-Feb-1996 #text_change 23-Mar-2001
C;Accession: S52185
R;Ellis, J.S.
submitted to the EMBL Data Library, October 1994
A;Description: Antigenic and genetic variation in the haemagglutinin of recently circula
A;Reference number: S52173
A;Accession: S52185
A;Molecule type: mRNA
A;Residues: 1-347 <ELL>
A;Cross-references: EMBL:Z46403; NID:g609043; PIDN:CAA86538.1; PID:g609044
A;Experimental source: isolate A/England/67/94
C;Superfamily: influenza virus hemagglutinin

Query Match 68.8%; Score 55; DB 2; Length 347;
Best Local Similarity 100.0%; Pred. No. 0.15;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 YPYDVPDYA 9
||||||||| |
Db 98 YPYDVPDYA 106

RESULT 15
S52186
hemagglutinin - influenza A virus (isolate A/England/68/94) (fragment)
C;Species: influenza A virus
A;Variety: isolate A/England/68/94
C;Date: 07-May-1995 #sequence_revision 23-Feb-1996 #text_change 23-Mar-2001
C;Accession: S52186
R;Ellis, J.S.
submitted to the EMBL Data Library, October 1994
A;Description: Antigenic and genetic variation in the haemagglutinin of recently circ
A;Reference number: S52173
A;Accession: S52186
A;Molecule type: mRNA
A;Residues: 1-347 <ELL>
A;Cross-references: EMBL:Z46404; NID:g609045; PIDN:CAA86539.1; PID:g609046
A;Experimental source: isolate A/England/68/94
C;Superfamily: influenza virus hemagglutinin

Query Match 68.8%; Score 55; DB 2; Length 347;
Best Local Similarity 100.0%; Pred. No. 0.15;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 YPYDVPDYA 9
||||||||| |
Db 98 YPYDVPDYA 106

Search completed: April 11, 2002, 10:16:16
Job time: 126 sec

GenCore version 4.5

OM protein - protein search, using sw model

Run on: April 11, 2002, 10:22:00 ; Search time 18.1 Seconds
(without alignments)
28.360 Million cell updates/sec

Title: US-09-284-787-2
Perfect score: 80
Sequence: 1 YPYDVPDYAGSGSK 14

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 100059 seqs, 36664827 residues

Total number of hits satisfying chosen parameters: 100059

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	55	68.8	249	1	HEMA_IAX31	P03438 influenza a
2	55	68.8	328	1	HEMA_IAEN6	P04664 influenza a
3	55	68.8	328	1	HEMA_IAQU7	P04663 influenza a
4	55	68.8	550	1	HEMA_IABAN	P03441 influenza a
5	55	68.8	550	1	HEMA_IADH1	P12582 influenza a
6	55	68.8	550	1	HEMA_IADH2	P12583 influenza a
7	55	68.8	550	1	HEMA_IADH3	P12584 influenza a
8	55	68.8	550	1	HEMA_IADH4	P12585 influenza a
9	55	68.8	550	1	HEMA_IADH6	P12587 influenza a
10	55	68.8	550	1	HEMA_IADH7	P12588 influenza a
11	55	68.8	550	1	HEMA_IADHK	P43257 influenza a
12	55	68.8	550	1	HEMA_IADHL	P43258 influenza a
13	55	68.8	550	1	HEMA_IADHM	P43259 influenza a
14	55	68.8	550	1	HEMA_IAGHK	P43260 influenza a
15	55	68.8	550	1	HEMA_IAME6	P12589 influenza a
16	55	68.8	550	1	HEMA_IAZH2	P11133 influenza a
17	55	68.8	550	1	HEMA_IAZH3	P11134 influenza a
18	55	68.8	565	1	HEMA_IAHAL	P16994 influenza a
19	55	68.8	565	1	HEMA_IAHMI	P15658 influenza a
20	55	68.8	566	1	HEMA_IAAIC	P03437 influenza a
21	55	68.8	566	1	HEMA_IADM2	P26135 influenza a
22	55	68.8	566	1	HEMA_IAEN7	P03440 influenza a
23	55	68.8	566	1	HEMA_IAMAO	P26138 influenza a
24	55	68.8	566	1	HEMA_IAME1	P03449 influenza a
25	55	68.8	566	1	HEMA_IAME2	P03439 influenza a
26	55	68.8	566	1	HEMA_IANT6	P03436 influenza a
27	55	68.8	566	1	HEMA_IAUDO	P19106 influenza a
28	55	68.8	566	1	HEMA_IAZCO	P26139 influenza a
29	55	68.8	567	1	HEMA_IAVI7	P03435 influenza a
30	54	67.5	565	1	HEMA_IAHFO	P16995 influenza a
31	54	67.5	565	1	HEMA_IAHK6	P19699 influenza a
32	54	67.5	565	1	HEMA_IAHK7	P16996 influenza a
33	54	67.5	565	1	HEMA_IAHNM	P16997 influenza a
34	54	67.5	565	1	HEMA_IAHRO	P16998 influenza a
35	54	67.5	565	1	HEMA_IAHSA	P16999 influenza a
36	54	67.5	565	1	HEMA_IAHSU	Q08011 influenza a
37	54	67.5	565	1	HEMA_IAHTE	P17001 influenza a
38	54	67.5	565	1	HEMA_IAHUR	P17002 influenza a
39	54	67.5	566	1	HEMA_IADU3	P03442 influenza a
40	52.5	65.6	566	1	HEMA_IADA3	P26134 influenza a
41	52.5	65.6	566	1	HEMA_IAZUK	P26141 influenza a
42	51.5	64.4	563	1	HEMA_IAMAA	P26137 influenza a
43	51.5	64.4	568	1	HEMA_IAMAB	P26136 influenza a
44	50	62.5	565	1	HEMA_IAHTO	P17000 influenza a
45	47	58.8	550	1	HEMA_IADH5	P12586 influenza a

ALIGNMENTS

RESULT 1
HEMA_IAX31
ID HEMA_IAX31 STANDARD; PRT; 249 AA.
AC P03438;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE HEMAGGLUTININ PRECURSOR [CONTAINS: HEMAGGLUTININ HA1 CHAIN]
DE (FRAGMENT).
GN HA.
OS Influenza A virus (strain X-31).
OC Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;
OC Influenza virus A and B group; Influenza A viruses;
OC Influenza A virus.
OX NCBI_TaxID=11489;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=81030852; PubMed=7421990;
RA Gething M.-J., Bye J., Skehel J.J., Waterfield M.;
RT "Cloning and DNA sequence of double-stranded copies of haemagglutinin
RT genes from H2 and H3 strains elucidates antigenic shift and drift in
RT human influenza virus.";
RL Nature 287:301-306(1980).
CC -!- FUNCTION: HEMAGGLUTININ IS RESPONSIBLE FOR ATTACHING THE VIRUS TO
CC CELL RECEPTORS AND FOR INITIATING INFECTION.
CC -!- SUBUNIT: HOMOTRIMER. EACH OF THE MONOMERS IS FORMED BY TWO CHAINS
CC (HA1 AND HA2) LINKED BY A DISULFIDE BOND.
CC -!- SIMILARITY: BELONGS TO THE INFLUENZA HEMAGGLUTININ FAMILY.
DR PIR; A04051; HMIVH.
DR PIR; A93233; A93233.
DR HSSP; P03437; 5HMG.
DR InterPro; IPR001364; Hemagglutn.
DR Pfam; PF00509; Hemagglutinin; 1.
DR ProDom; PD000225; Hemagglutn; 1.
KW Envelope protein; Hemagglutinin; Glycoprotein; Signal.
FT SIGNAL 1 16
FT CHAIN 17 >249 HEMAGGLUTININ HA1 CHAIN.
FT CARBOHYD 24 24 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 38 38 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 54 54 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 97 97 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 181 181 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT NON_TER 249 249
SQ SEQUENCE 249 AA; 27373 MW; 9537AA970BB79183 CRC64;

Query Match 68.8%; Score 55; DB 1; Length 249;
Best Local Similarity 100.0%; Pred. No. 0.052;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 YPYDVPDYA 9
|||||||||
Db 114 YPYDVPDYA 122

RESULT 2

```
HEMA_IAEN6
ID   HEMA_IAEN6     STANDARD;      PRT;   328 AA.
AC   P04664;
DT   13-AUG-1987 (Rel. 05, Created)
DT   13-AUG-1987 (Rel. 05, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   HEMAGGLUTININ [CONTAINS: HEMAGGLUTININ HA1 CHAIN] (FRAGMENT).
GN   HA.
OS   Influenza A virus (strain A/England/878/69).
OC   Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;
OC   Influenza virus A and B group; Influenza A viruses;
OC   Influenza A virus.
OX   NCBI_TaxID=11377;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=81194918; PubMed=6164798;
RA   Sleigh M.J., Both G.W., Underwood P.A., Bender V.J.;
RT   "Antigenic drift in the hemagglutinin of the Hong Kong influenza
RT   subtype: correlation of amino acid changes with alterations in viral
RT   antigenicity.";
RL   J. Virol. 37:845-853(1981).
CC   -!- FUNCTION: HEMAGGLUTININ IS RESPONSIBLE FOR ATTACHING THE VIRUS TO
CC       CELL RECEPTORS AND FOR INITIATING INFECTION.
CC   -!- SUBUNIT: HOMOTRIMER. EACH OF THE MONOMERS IS FORMED BY TWO CHAINS
CC       (HA1 AND HA2) LINKED BY A DISULFIDE BOND.
CC   -!- SIMILARITY: BELONGS TO THE INFLUENZA HEMAGGLUTININ FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; K03335; AAA43184.1; -.
DR   InterPro; IPR001364; Hemagglutn.
DR   Pfam; PF00509; Hemagglutinin; 1.
DR   ProDom; PD000225; Hemagglutn; 1.
KW   Envelope protein; Hemagglutinin; Glycoprotein.
FT   NON_TER       1      1
FT   CHAIN         1    328       HEMAGGLUTININ HA1 CHAIN.
FT   CARBOHYD      8      8       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     22     22       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     38     38       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     63     63       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    165    165       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    285    285       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   NON_TER     328    328
SQ   SEQUENCE   328 AA;  36072 MW;  9C3A86B3A8D856FE CRC64;


  Query Match             68.8%;  Score 55;  DB 1;  Length 328;
  Best Local Similarity   100.0%;  Pred. No. 0.069;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 YPYDVPDYA 9
           |||||||||
Db      98 YPYDVPDYA 106


RESULT   3
HEMA_IAQU7
ID   HEMA_IAQU7     STANDARD;      PRT;   328 AA.
AC   P04663;
DT   13-AUG-1987 (Rel. 05, Created)
DT   13-AUG-1987 (Rel. 05, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   HEMAGGLUTININ [CONTAINS: HEMAGGLUTININ HA1 CHAIN] (FRAGMENT).
GN   HA.
OS   Influenza A virus (strain A/Qu/7/70).
OC   Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;
OC   Influenza virus A and B group; Influenza A viruses;
OC   Influenza A virus.
OX   NCBI_TaxID=11456;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=81194918; PubMed=6164798;
RA   Sleigh M.J., Both G.W., Underwood P.A., Bender V.J.;
RT   "Antigenic drift in the hemagglutinin of the Hong Kong influenza
RT   subtype: correlation of amino acid changes with alterations in viral
RT   antigenicity.";
RL   J. Virol. 37:845-853(1981).
CC   -!- FUNCTION: HEMAGGLUTININ IS RESPONSIBLE FOR ATTACHING THE VIRUS TO
CC       CELL RECEPTORS AND FOR INITIATING INFECTION.
CC   -!- SUBUNIT: HOMOTRIMER. EACH OF THE MONOMERS IS FORMED BY TWO CHAINS
CC       (HA1 AND HA2) LINKED BY A DISULFIDE BOND.
CC   -!- SIMILARITY: BELONGS TO THE INFLUENZA HEMAGGLUTININ FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; K03338; AAA43195.1; -.
DR   InterPro; IPR001364; Hemagglutn.
DR   Pfam; PF00509; Hemagglutinin; 1.
DR   ProDom; PD000225; Hemagglutn; 1.
KW   Envelope protein; Hemagglutinin; Glycoprotein.
FT   NON_TER       1      1
FT   CHAIN         1    328       HEMAGGLUTININ HA1 CHAIN.
FT   CARBOHYD      8      8       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     22     22       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     38     38       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     81     81       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    165    165       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    285    285       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   NON_TER     328    328
SQ   SEQUENCE   328 AA;  36059 MW;  E30A962FD6DD805B CRC64;


  Query Match             68.8%;  Score 55;  DB 1;  Length 328;
  Best Local Similarity   100.0%;  Pred. No. 0.069;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 YPYDVPDYA 9
           |||||||||
Db      98 YPYDVPDYA 106


RESULT   4
HEMA_IABAN
ID   HEMA_IABAN     STANDARD;      PRT;   550 AA.
AC   P03441; Q83961; Q83962;
DT   21-JUL-1986 (Rel. 01, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   HEMAGGLUTININ PRECURSOR [CONTAINS: HEMAGGLUTININ HA1 CHAIN;
DE   HEMAGGLUTININ HA2 CHAIN] (FRAGMENT).
GN   HA.
OS   Influenza A virus (strain A/Bangkok/1/79).
OC   Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;
OC   Influenza virus A and B group; Influenza A viruses;
OC   Influenza A virus.
OX   NCBI_TaxID=11325;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=82033259; PubMed=6169840;
RA   Both G.W., Sleigh M.J.;
RT   "Conservation and variation in the hemagglutinins of Hong Kong
RT   subtype influenza viruses during antigenic drift.";
RL   J. Virol. 39:663-672(1981).
CC   -!- FUNCTION: HEMAGGLUTININ IS RESPONSIBLE FOR ATTACHING THE VIRUS TO
```

```
CC       CELL RECEPTORS AND FOR INITIATING INFECTION.
CC   -!- SUBUNIT: HOMOTRIMER. EACH OF THE MONOMERS IS FORMED BY TWO CHAINS
CC       (HA1 AND HA2) LINKED BY A DISULFIDE BOND.
CC   -!- SIMILARITY: BELONGS TO THE INFLUENZA HEMAGGLUTININ FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; J02092; AAA43182.1; -.
DR   PIR; A04052; HMIV6.
DR   InterPro; IPR001364; Hemagglutn.
DR   Pfam; PF00509; Hemagglutinin; 1.
DR   ProDom; PD000225; Hemagglutn; 1.
KW   Envelope protein; Hemagglutinin; Glycoprotein.
FT   NON_TER       1      1
FT   CHAIN         1    328       HEMAGGLUTININ HA1 CHAIN.
FT   CHAIN       330    550       HEMAGGLUTININ HA2 CHAIN.
FT   CARBOHYD      8      8       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     22     22       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     38     38       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     63     63       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    126    126       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    165    165       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    285    285       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    483    483       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   550 AA;  61699 MW;  FBD2EC200689CBE5 CRC64;


  Query Match             68.8%;  Score 55;  DB 1;  Length 550;
  Best Local Similarity   100.0%;  Pred. No. 0.12;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 YPYDVPDYA 9
           |||||||||
Db      98 YPYDVPDYA 106


RESULT   5
HEMA_IADH1
ID   HEMA_IADH1     STANDARD;      PRT;   550 AA.
AC   P12582; Q84021; Q84022;
DT   01-OCT-1989 (Rel. 12, Created)
DT   01-OCT-1989 (Rel. 12, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   HEMAGGLUTININ PRECURSOR [CONTAINS: HEMAGGLUTININ HA1 CHAIN;
DE   HEMAGGLUTININ HA2 CHAIN] (FRAGMENT).
GN   HA.
OS   Influenza A virus (strain A/Duck/Hokkaido/5/77).
OC   Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;
OC   Influenza virus A and B group; Influenza A viruses;
OC   Influenza A virus.
OX   NCBI_TaxID=11357;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=87265458; PubMed=2440178;
RA   Kida H., Kawaoka Y., Naeve C.W., Webster R.G.;
RT   "Antigenic and genetic conservation of H3 influenza virus in wild
RT   ducks.";
RL   Virology 159:109-119(1987).
CC   -!- FUNCTION: HEMAGGLUTININ IS RESPONSIBLE FOR ATTACHING THE VIRUS TO
CC       CELL RECEPTORS AND FOR INITIATING INFECTION.
CC   -!- SUBUNIT: HOMOTRIMER. EACH OF THE MONOMERS IS FORMED BY TWO CHAINS
CC       (HA1 AND HA2) LINKED BY A DISULFIDE BOND.
CC   -!- SIMILARITY: BELONGS TO THE INFLUENZA HEMAGGLUTININ FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M16737; AAA43143.1; -.
DR   PIR; A27813; HMIV77.
DR   InterPro; IPR001364; Hemagglutn.
DR   Pfam; PF00509; Hemagglutinin; 1.
DR   ProDom; PD000225; Hemagglutn; 1.
KW   Envelope protein; Hemagglutinin; Glycoprotein.
FT   NON_TER       1      1
FT   CHAIN         1    328       HEMAGGLUTININ HA1 CHAIN.
FT   CHAIN       330    550       HEMAGGLUTININ HA2 CHAIN.
FT   CARBOHYD      8      8       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     22     22       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     38     38       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    165    165       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    285    285       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    483    483       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   550 AA;  61705 MW;  7E7ACFE716FC969A CRC64;


  Query Match             68.8%;  Score 55;  DB 1;  Length 550;
  Best Local Similarity   100.0%;  Pred. No. 0.12;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 YPYDVPDYA 9
           |||||||||
Db      98 YPYDVPDYA 106


RESULT   6
HEMA_IADH2
ID   HEMA_IADH2     STANDARD;      PRT;   550 AA.
AC   P12583; Q84011;
DT   01-OCT-1989 (Rel. 12, Created)
DT   01-APR-1990 (Rel. 14, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   HEMAGGLUTININ PRECURSOR [CONTAINS: HEMAGGLUTININ HA1 CHAIN;
DE   HEMAGGLUTININ HA2 CHAIN] (FRAGMENT).
GN   HA.
OS   Influenza A virus (strain A/Duck/Hokkaido/8/80).
OC   Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;
OC   Influenza virus A and B group; Influenza A viruses;
OC   Influenza A virus.
OX   NCBI_TaxID=11358;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=87265458; PubMed=2440178;
RA   Kida H., Kawaoka Y., Naeve C.W., Webster R.G.;
RT   "Antigenic and genetic conservation of H3 influenza virus in wild
RT   ducks.";
RL   Virology 159:109-119(1987).
CC   -!- FUNCTION: HEMAGGLUTININ IS RESPONSIBLE FOR ATTACHING THE VIRUS TO
CC       CELL RECEPTORS AND FOR INITIATING INFECTION.
CC   -!- SUBUNIT: HOMOTRIMER. EACH OF THE MONOMERS IS FORMED BY TWO CHAINS
CC       (HA1 AND HA2) LINKED BY A DISULFIDE BOND.
CC   -!- SIMILARITY: BELONGS TO THE INFLUENZA HEMAGGLUTININ FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M16738; AAA43144.1; -.
DR   PIR; B27813; HMIV80.
DR   InterPro; IPR001364; Hemagglutn.
DR   Pfam; PF00509; Hemagglutinin; 1.
DR   ProDom; PD000225; Hemagglutn; 1.
```

```
KW   Envelope protein; Hemagglutinin; Glycoprotein.
FT   NON_TER       1      1
FT   CHAIN         1    328       HEMAGGLUTININ HA1 CHAIN.
FT   CHAIN       330    550       HEMAGGLUTININ HA2 CHAIN.
FT   CARBOHYD      8      8       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     22     22       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     38     38       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    165    165       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    285    285       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    483    483       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CONFLICT    137    137       K -> N (IN PIR DATA BANK).
SQ   SEQUENCE   550 AA;  61659 MW;  A107023ACC9CC353 CRC64;


  Query Match             68.8%;  Score 55;  DB 1;  Length 550;
  Best Local Similarity   100.0%;  Pred. No. 0.12;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 YPYDVPDYA 9
           |||||||||
Db      98 YPYDVPDYA 106


RESULT   7
HEMA_IADH3
ID   HEMA_IADH3     STANDARD;      PRT;   550 AA.
AC   P12584; Q84012; Q89793;
DT   01-OCT-1989 (Rel. 12, Created)
DT   01-OCT-1989 (Rel. 12, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   HEMAGGLUTININ PRECURSOR [CONTAINS: HEMAGGLUTININ HA1 CHAIN;
DE   HEMAGGLUTININ HA2 CHAIN] (FRAGMENT).
GN   HA.
OS   Influenza A virus (strain A/Duck/Hokkaido/33/80).
OC   Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;
OC   Influenza virus A and B group; Influenza A viruses;
OC   Influenza A virus.
OX   NCBI_TaxID=11359;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=87265458; PubMed=2440178;
RA   Kida H., Kawaoka Y., Naeve C.W., Webster R.G.;
RT   "Antigenic and genetic conservation of H3 influenza virus in wild
RT   ducks.";
RL   Virology 159:109-119(1987).
CC   -!- FUNCTION: HEMAGGLUTININ IS RESPONSIBLE FOR ATTACHING THE VIRUS TO
CC       CELL RECEPTORS AND FOR INITIATING INFECTION.
CC   -!- SUBUNIT: HOMOTRIMER. EACH OF THE MONOMERS IS FORMED BY TWO CHAINS
CC       (HA1 AND HA2) LINKED BY A DISULFIDE BOND.
CC   -!- SIMILARITY: BELONGS TO THE INFLUENZA HEMAGGLUTININ FAMILY.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; M16739; AAA43145.1; -.
DR   PIR; C27813; HMIV33.
DR   InterPro; IPR001364; Hemagglutn.
DR   Pfam; PF00509; Hemagglutinin; 1.
DR   ProDom; PD000225; Hemagglutn; 1.
KW   Envelope protein; Hemagglutinin; Glycoprotein.
FT   NON_TER       1      1
FT   CHAIN         1    328       HEMAGGLUTININ HA1 CHAIN.
FT   CHAIN       330    550       HEMAGGLUTININ HA2 CHAIN.
FT   CARBOHYD      8      8       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     22     22       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     38     38       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    165    165       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    285    285       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    483    483       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   550 AA;  61577 MW;  6C30BF67CFDCB7DE CRC64;


  Query Match             68.8%;  Score 55;  DB 1;  Length 550;
  Best Local Similarity   100.0%;  Pred. No. 0.12;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 YPYDVPDYA 9
           |||||||||
Db      98 YPYDVPDYA 106


RESULT   8
HEMA_IADH4
ID   HEMA_IADH4     STANDARD;      PRT;   550 AA.
AC   P12585; Q84013; Q84014;
DT   01-OCT-1989 (Rel. 12, Created)
DT   01-OCT-1989 (Rel. 12, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   HEMAGGLUTININ PRECURSOR [CONTAINS: HEMAGGLUTININ HA1 CHAIN;
DE   HEMAGGLUTININ HA2 CHAIN] (FRAGMENT).
GN   HA.
OS   Influenza A virus (strain A/Duck/Hokkaido/7/82).
OC   Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;
OC   Influenza virus A and B group; Influenza A viruses;
OC   Influenza A virus.
OX   NCBI_TaxID=11360;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=87265458; PubMed=2440178;
RA   Kida H., Kawaoka Y., Naeve C.W., Webster R.G.;
RT   "Antigenic and genetic conservation of H3 influenza virus in wild
RT   ducks.";
RL   Virology 159:109-119(1987).
CC   -!- FUNCTION: HEMAGGLUTININ IS RESPONSIBLE FOR ATTACHING THE VIRUS TO
CC       CELL RECEPTORS AND FOR INITIATING INFECTION.
CC   -!- SUBUNIT: HOMOTRIMER. EACH OF THE MONOMERS IS FORMED BY TWO CHAINS
CC       (HA1 AND HA2) LINKED BY A DISULFIDE BOND.
CC   -!- SIMILARITY: BELONGS TO THE INFLUENZA HEMAGGLUTININ FAMILY.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; M16740; AAA43146.1; -.
DR   PIR; D27813; HMIV89.
DR   InterPro; IPR001364; Hemagglutn.
DR   Pfam; PF00509; Hemagglutinin; 1.
DR   ProDom; PD000225; Hemagglutn; 1.
KW   Envelope protein; Hemagglutinin; Glycoprotein.
FT   NON_TER       1      1
FT   CHAIN         1    328       HEMAGGLUTININ HA1 CHAIN.
FT   CHAIN       330    550       HEMAGGLUTININ HA2 CHAIN.
FT   CARBOHYD      8      8       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     22     22       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     38     38       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    165    165       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    285    285       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    483    483       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   550 AA;  61664 MW;  A16B2CF8CBBBD9D0 CRC64;


  Query Match             68.8%;  Score 55;  DB 1;  Length 550;
  Best Local Similarity   100.0%;  Pred. No. 0.12;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 YPYDVPDYA 9
           |||||||||
```

```
Db     98 YPYDVPDYA 106


RESULT   9
HEMA_IADH6
ID   HEMA_IADH6     STANDARD;      PRT;   550 AA.
AC   P12587; Q84017;
DT   01-OCT-1989 (Rel. 12, Created)
DT   01-APR-1990 (Rel. 14, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   HEMAGGLUTININ PRECURSOR [CONTAINS: HEMAGGLUTININ HA1 CHAIN;
DE   HEMAGGLUTININ HA2 CHAIN] (FRAGMENT).
GN   HA.
OS   Influenza A virus (strain A/Duck/Hokkaido/9/85).
OC   Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;
OC   Influenza virus A and B group; Influenza A viruses;
OC   Influenza A virus.
OX   NCBI_TaxID=11362;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=87265458; PubMed=2440178;
RA   Kida H., Kawaoka Y., Naeve C.W., Webster R.G.;
RT   "Antigenic and genetic conservation of H3 influenza virus in wild
RT   ducks.";
RL   Virology 159:109-119(1987).
CC   -!- FUNCTION: HEMAGGLUTININ IS RESPONSIBLE FOR ATTACHING THE VIRUS TO
CC       CELL RECEPTORS AND FOR INITIATING INFECTION.
CC   -!- SUBUNIT: HOMOTRIMER. EACH OF THE MONOMERS IS FORMED BY TWO CHAINS
CC       (HA1 AND HA2) LINKED BY A DISULFIDE BOND.
CC   -!- SIMILARITY: BELONGS TO THE INFLUENZA HEMAGGLUTININ FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M16742; AAA43148.1; -.
DR   PIR; F27813; HMIV98.
DR   InterPro; IPR001364; Hemagglutn.
DR   Pfam; PF00509; Hemagglutinin; 1.
DR   ProDom; PD000225; Hemagglutn; 1.
KW   Envelope protein; Hemagglutinin; Glycoprotein.
FT   NON_TER       1      1
FT   CHAIN         1    328       HEMAGGLUTININ HA1 CHAIN.
FT   CHAIN       330    550       HEMAGGLUTININ HA2 CHAIN.
FT   CARBOHYD     22     22       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     38     38       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    165    165       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    285    285       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    483    483       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CONFLICT      8      8       Y -> N (IN PIR DATA BANK).
SQ   SEQUENCE   550 AA;  61711 MW;  67BCD85F44736CFE CRC64;


  Query Match             68.8%;  Score 55;  DB 1;  Length 550;
  Best Local Similarity   100.0%;  Pred. No. 0.12;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 YPYDVPDYA 9
           |||||||||
Db      98 YPYDVPDYA 106


RESULT  10
HEMA_IADH7
ID   HEMA_IADH7     STANDARD;      PRT;   550 AA.
AC   P12588; Q84018; Q89470;
DT   01-OCT-1989 (Rel. 12, Created)
DT   01-OCT-1989 (Rel. 12, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   HEMAGGLUTININ PRECURSOR [CONTAINS: HEMAGGLUTININ HA1 CHAIN;
DE   HEMAGGLUTININ HA2 CHAIN] (FRAGMENT).
GN   HA.
OS   Influenza A virus (strain A/Duck/Hokkaido/10/85).
OC   Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;
OC   Influenza virus A and B group; Influenza A viruses;
OC   Influenza A virus.
OX   NCBI_TaxID=11363;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=87265458; PubMed=2440178;
RA   Kida H., Kawaoka Y., Naeve C.W., Webster R.G.;
RT   "Antigenic and genetic conservation of H3 influenza virus in wild
RT   ducks.";
RL   Virology 159:109-119(1987).
CC   -!- FUNCTION: HEMAGGLUTININ IS RESPONSIBLE FOR ATTACHING THE VIRUS TO
CC       CELL RECEPTORS AND FOR INITIATING INFECTION.
CC   -!- SUBUNIT: HOMOTRIMER. EACH OF THE MONOMERS IS FORMED BY TWO CHAINS
CC       (HA1 AND HA2) LINKED BY A DISULFIDE BOND.
CC   -!- SIMILARITY: BELONGS TO THE INFLUENZA HEMAGGLUTININ FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M16743; AAA43149.1; -.
DR   PIR; G27813; HMIV15.
DR   InterPro; IPR001364; Hemagglutn.
DR   Pfam; PF00509; Hemagglutinin; 1.
DR   ProDom; PD000225; Hemagglutn; 1.
KW   Envelope protein; Hemagglutinin; Glycoprotein.
FT   NON_TER       1      1
FT   CHAIN         1    328       HEMAGGLUTININ HA1 CHAIN.
FT   CHAIN       330    550       HEMAGGLUTININ HA2 CHAIN.
FT   CARBOHYD      8      8       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     22     22       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     38     38       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    165    165       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    285    285       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    483    483       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   550 AA;  61761 MW;  6EF81793281D53EB CRC64;


  Query Match             68.8%;  Score 55;  DB 1;  Length 550;
  Best Local Similarity   100.0%;  Pred. No. 0.12;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 YPYDVPDYA 9
           |||||||||
Db      98 YPYDVPDYA 106


RESULT  11
HEMA_IADHK
ID   HEMA_IADHK     STANDARD;      PRT;   550 AA.
AC   P43257;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   HEMAGGLUTININ PRECURSOR [CONTAINS: HEMAGGLUTININ HA1 CHAIN;
DE   HEMAGGLUTININ HA2 CHAIN] (FRAGMENT).
GN   HA.
OS   Influenza A virus (strain A/Duck/Hong Kong/7/75).
OC   Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;
OC   Influenza virus A and B group; Influenza A viruses;
OC   Influenza A virus.
OX   NCBI_TaxID=11364;
RN   [1]
RP   SEQUENCE FROM N.A.
```

```
RX   MEDLINE=91341491; PubMed=1875195;
RA   Yasuda J., Shortridge K.F., Shimizu Y., Kida H.;
RT   "Molecular evidence for a role of domestic ducks in the introduction
RT   of avian H3 influenza viruses to pigs in southern China, where the
RT   A/Hong Kong/68 (H3N2) strain emerged.";
RL   J. Gen. Virol. 72:2007-2010(1991).
CC   -!- FUNCTION: HEMAGGLUTININ IS RESPONSIBLE FOR ATTACHING THE VIRUS TO
CC       CELL RECEPTORS AND FOR INITIATING INFECTION.
CC   -!- SUBUNIT: HOMOTRIMER. EACH OF THE MONOMERS IS FORMED BY TWO CHAINS
CC       (HA1 AND HA2) LINKED BY A DISULFIDE BOND.
CC   -!- SIMILARITY: BELONGS TO THE INFLUENZA HEMAGGLUTININ FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; D00929; BAA00769.1; -.
DR   HSSP; P03437; 5HMG.
DR   InterPro; IPR001364; Hemagglutn.
DR   Pfam; PF00509; Hemagglutinin; 1.
DR   ProDom; PD000225; Hemagglutn; 1.
KW   Envelope protein; Hemagglutinin; Glycoprotein.
FT   NON_TER       1      1
FT   CHAIN         1    328       HEMAGGLUTININ HA1 CHAIN.
FT   CHAIN       330    550       HEMAGGLUTININ HA2 CHAIN.
FT   CARBOHYD      8      8       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     22     22       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     38     38       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    165    165       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    285    285       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    483    483       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   550 AA;  61549 MW;  864639B829FE1BA9 CRC64;


  Query Match             68.8%;  Score 55;  DB 1;  Length 550;
  Best Local Similarity   100.0%;  Pred. No. 0.12;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 YPYDVPDYA 9
           |||||||||
Db      98 YPYDVPDYA 106


RESULT  12
HEMA_IADHL
ID   HEMA_IADHL     STANDARD;      PRT;   550 AA.
AC   P43258;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   HEMAGGLUTININ PRECURSOR [CONTAINS: HEMAGGLUTININ HA1 CHAIN;
DE   HEMAGGLUTININ HA2 CHAIN] (FRAGMENT).
GN   HA.
OS   Influenza A virus (strain A/Duck/Hong Kong/64/76).
OC   Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;
OC   Influenza virus A and B group; Influenza A viruses;
OC   Influenza A virus.
OX   NCBI_TaxID=45412;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=91341491; PubMed=1875195;
RA   Yasuda J., Shortridge K.F., Shimizu Y., Kida H.;
RT   "Molecular evidence for a role of domestic ducks in the introduction
RT   of avian H3 influenza viruses to pigs in southern China, where the
RT   A/Hong Kong/68 (H3N2) strain emerged.";
RL   J. Gen. Virol. 72:2007-2010(1991).
CC   -!- FUNCTION: HEMAGGLUTININ IS RESPONSIBLE FOR ATTACHING THE VIRUS TO
CC       CELL RECEPTORS AND FOR INITIATING INFECTION.
CC   -!- SUBUNIT: HOMOTRIMER. EACH OF THE MONOMERS IS FORMED BY TWO CHAINS
CC       (HA1 AND HA2) LINKED BY A DISULFIDE BOND.
CC   -!- SIMILARITY: BELONGS TO THE INFLUENZA HEMAGGLUTININ FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; D00931; BAA00771.1; -.
DR   HSSP; P03437; 2HMG.
DR   InterPro; IPR001364; Hemagglutn.
DR   Pfam; PF00509; Hemagglutinin; 1.
DR   ProDom; PD000225; Hemagglutn; 1.
KW   Envelope protein; Hemagglutinin; Glycoprotein.
FT   NON_TER       1      1
FT   CHAIN         1    328       HEMAGGLUTININ HA1 CHAIN.
FT   CHAIN       330    550       HEMAGGLUTININ HA2 CHAIN.
FT   CARBOHYD      8      8       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     22     22       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     38     38       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    165    165       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    285    285       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    483    483       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   550 AA;  61718 MW;  A351C56789E4BE9A CRC64;


  Query Match             68.8%;  Score 55;  DB 1;  Length 550;
  Best Local Similarity   100.0%;  Pred. No. 0.12;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 YPYDVPDYA 9
           |||||||||
Db      98 YPYDVPDYA 106


RESULT  13
HEMA_IADHM
ID   HEMA_IADHM     STANDARD;      PRT;   550 AA.
AC   P43259;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   HEMAGGLUTININ PRECURSOR [CONTAINS: HEMAGGLUTININ HA1 CHAIN;
DE   HEMAGGLUTININ HA2 CHAIN] (FRAGMENT).
GN   HA.
OS   Influenza A virus (strain A/Duck/Hong Kong/231/77).
OC   Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;
OC   Influenza virus A and B group; Influenza A viruses;
OC   Influenza A virus.
OX   NCBI_TaxID=45411;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=91341491; PubMed=1875195;
RA   Yasuda J., Shortridge K.F., Shimizu Y., Kida H.;
RT   "Molecular evidence for a role of domestic ducks in the introduction
RT   of avian H3 influenza viruses to pigs in southern China, where the
RT   A/Hong Kong/68 (H3N2) strain emerged.";
RL   J. Gen. Virol. 72:2007-2010(1991).
CC   -!- FUNCTION: HEMAGGLUTININ IS RESPONSIBLE FOR ATTACHING THE VIRUS TO
CC       CELL RECEPTORS AND FOR INITIATING INFECTION.
CC   -!- SUBUNIT: HOMOTRIMER. EACH OF THE MONOMERS IS FORMED BY TWO CHAINS
CC       (HA1 AND HA2) LINKED BY A DISULFIDE BOND.
CC   -!- SIMILARITY: BELONGS TO THE INFLUENZA HEMAGGLUTININ FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
```

```
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; D00932; BAA00772.1; -.
DR   InterPro; IPR001364; Hemagglutn.
DR   Pfam; PF00509; Hemagglutinin; 1.
DR   ProDom; PD000225; Hemagglutn; 1.
KW   Envelope protein; Hemagglutinin; Glycoprotein.
FT   NON_TER       1      1
FT   CHAIN         1    328       HEMAGGLUTININ HA1 CHAIN.
FT   CHAIN       330    550       HEMAGGLUTININ HA2 CHAIN.
FT   CARBOHYD     22     22       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     38     38       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    165    165       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    285    285       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    483    483       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   550 AA;  61762 MW;  6FEF8B49488C191A CRC64;


  Query Match             68.8%;  Score 55;  DB 1;  Length 550;
  Best Local Similarity   100.0%;  Pred. No. 0.12;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 YPYDVPDYA 9
          |||||||||
Db     98 YPYDVPDYA 106


RESULT  14
HEMA_IAGHK
ID   HEMA_IAGHK     STANDARD;      PRT;   550 AA.
AC   P43260;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update).
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   HEMAGGLUTININ PRECURSOR [CONTAINS: HEMAGGLUTININ HA1 CHAIN;
DE   HEMAGGLUTININ HA2 CHAIN] (FRAGMENT).
GN   HA.
OS   Influenza A virus (strain A/Goose/Hong Kong/10/76).
OC   Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;
OC   Influenza virus A and B group; Influenza A viruses;
OC   Influenza A virus.
OX   NCBI_TaxID=45414;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=91341491; PubMed=1875195;
RA   Yasuda J., Shortridge K.F., Shimizu Y., Kida H.;
RT   "Molecular evidence for a role of domestic ducks in the introduction
RT   of avian H3 influenza viruses to pigs in southern China, where the
RT   A/Hong Kong/68 (H3N2) strain emerged.";
RL   J. Gen. Virol. 72:2007-2010(1991).
CC   -!- FUNCTION: HEMAGGLUTININ IS RESPONSIBLE FOR ATTACHING THE VIRUS TO
CC       CELL RECEPTORS AND FOR INITIATING INFECTION.
CC   -!- SUBUNIT: HOMOTRIMER. EACH OF THE MONOMERS IS FORMED BY TWO CHAINS
CC       (HA1 AND HA2) LINKED BY A DISULFIDE BOND.
CC   -!- SIMILARITY: BELONGS TO THE INFLUENZA HEMAGGLUTININ FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; D00930; BAA00770.1; -.
DR   HSSP; P03437; 5HMG.
DR   InterPro; IPR001364; Hemagglutn.
DR   Pfam; PF00509; Hemagglutinin; 1.
DR   ProDom; PD000225; Hemagglutn; 1.
KW   Envelope protein; Hemagglutinin; Glycoprotein.
FT   NON_TER       1      1
FT   CHAIN         1    328       HEMAGGLUTININ HA1 CHAIN.
FT   CHAIN       330    550       HEMAGGLUTININ HA2 CHAIN.
FT   CARBOHYD      8      8       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     22     22       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     38     38       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    165    165       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    285    285       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    483    483       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   550 AA;  61676 MW;  9A1E094DA28BACD2 CRC64;


  Query Match             68.8%;  Score 55;  DB 1;  Length 550;
  Best Local Similarity   100.0%;  Pred. No. 0.12;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 YPYDVPDYA 9
          |||||||||
Db     98 YPYDVPDYA 106


RESULT  15
HEMA_IAME6
ID   HEMA_IAME6     STANDARD;      PRT;   550 AA.
AC   P12589;
DT   01-OCT-1989 (Rel. 12, Created)
DT   01-OCT-1989 (Rel. 12, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   HEMAGGLUTININ PRECURSOR [CONTAINS: HEMAGGLUTININ HA1 CHAIN;
DE   HEMAGGLUTININ HA2 CHAIN] (FRAGMENT).
GN   HA.
OS   Influenza A virus (strain A/Memphis/6/86).
OC   Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;
OC   Influenza virus A and B group; Influenza A viruses;
OC   Influenza A virus.
OX   NCBI_TaxID=11440;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=88306236; PubMed=3407150;
RA   Katz J.M., Webster R.G.;
RT   "Antigenic and structural characterization of multiple subpopulations
RT   of H3N2 influenza virus from an individual.";
RL   Virology 165:446-456(1988).
CC   -!- FUNCTION: HEMAGGLUTININ IS RESPONSIBLE FOR ATTACHING THE VIRUS TO
CC       CELL RECEPTORS AND FOR INITIATING INFECTION.
CC   -!- SUBUNIT: HOMOTRIMER. EACH OF THE MONOMERS IS FORMED BY TWO CHAINS
CC       (HA1 AND HA2) LINKED BY A DISULFIDE BOND.
CC   -!- SIMILARITY: BELONGS TO THE INFLUENZA HEMAGGLUTININ FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M21648; AAA43275.1; -.
DR   PIR; A29245; HMIV86.
DR   HSSP; P03437; 2HMG.
DR   InterPro; IPR001364; Hemagglutn.
DR   Pfam; PF00509; Hemagglutinin; 1.
DR   ProDom; PD000225; Hemagglutn; 1.
KW   Envelope protein; Hemagglutinin; Glycoprotein.
FT   NON_TER       1      1
FT   CHAIN         1    328       HEMAGGLUTININ HA1 CHAIN.
FT   CHAIN       330    550       HEMAGGLUTININ HA2 CHAIN.
FT   CARBOHYD      8      8       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     22     22       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     38     38       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     63     63       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    126    126       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    165    165       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    246    246       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    285    285       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    483    483       N-LINKED (GLCNAC. . .) (POTENTIAL).
```

SQ SEQUENCE 550 AA; 61804 MW; 52C9F14B309310ED CRC64;

Query Match 68.8%; Score 55; DB 1; Length 550;
Best Local Similarity 100.0%; Pred. No. 0.12;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       1 YPYDVPDYA 9
           |||||||||
Db      98 YPYDVPDYA 106
```

Search completed: April 11, 2002, 10:22:00
Job time: 430 sec

GenCore version 4.5

OM protein - protein search, using sw model

Run on: April 11, 2002, 10:22:57 ; Search time 50.04 Seconds
(without alignments)
40.924 Million cell updates/sec

Title: US-09-284-787-2
Perfect score: 80
Sequence: 1 YPYDVPDYAGSGSK 14

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 473505 seqs, 146272329 residues

Total number of hits satisfying chosen parameters: 473505

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_17:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	55	68.8	236	12	Q9W8Z1	Q9w8z1 influenza a
2	55	68.8	236	12	Q9W8T9	Q9w8t9 influenza a
3	55	68.8	236	12	Q9W8J7	Q9w8j7 influenza a
4	55	68.8	286	12	Q9WI26	Q9wi26 influenza a
5	55	68.8	286	12	Q9WI25	Q9wi25 influenza a
6	55	68.8	286	12	Q9WI24	Q9wi24 influenza a
7	55	68.8	325	12	O40774	O40774 influenza a
8	55	68.8	325	12	Q9DL45	Q9dl45 influenza a
9	55	68.8	326	12	O40779	O40779 influenza a
10	55	68.8	326	12	O40780	O40780 influenza a
11	55	68.8	327	12	O92324	O92324 influenza a
12	55	68.8	327	12	O92325	O92325 influenza a
13	55	68.8	327	12	O92327	O92327 influenza a
14	55	68.8	327	12	Q9YT90	Q9yt90 influenza a
15	55	68.8	327	12	Q9YT89	Q9yt89 influenza a
16	55	68.8	327	12	Q9YT88	Q9yt88 influenza a
17	55	68.8	327	12	Q9YT87	Q9yt87 influenza a
18	55	68.8	327	12	Q9YT86	Q9yt86 influenza a
19	55	68.8	327	12	Q9YRU9	Q9yru9 influenza a
20	55	68.8	327	12	Q9YRU8	Q9yru8 influenza a
21	55	68.8	327	12	Q9YRU7	Q9yru7 influenza a
22	55	68.8	327	12	Q9YRU6	Q9yru6 influenza a
23	55	68.8	327	12	Q9YRU5	Q9yru5 influenza a
24	55	68.8	327	12	Q9YRU4	Q9yru4 influenza a
25	55	68.8	327	12	Q9YRU3	Q9yru3 influenza a
26	55	68.8	327	12	Q9YJ57	Q9yj57 influenza a
27	55	68.8	327	12	Q9YIN9	Q9yin9 influenza a
28	55	68.8	327	12	Q9Q884	Q9q884 influenza a
29	55	68.8	328	12	O40855	O40855 influenza a
30	55	68.8	328	12	O40867	O40867 influenza a
31	55	68.8	328	12	Q82567	Q82567 influenza a
32	55	68.8	328	12	Q82568	Q82568 influenza a
33	55	68.8	328	12	Q82569	Q82569 influenza a
34	55	68.8	328	12	Q82578	Q82578 influenza a
35	55	68.8	328	12	Q82579	Q82579 influenza a
36	55	68.8	328	12	Q82580	Q82580 influenza a
37	55	68.8	328	12	Q82581	Q82581 influenza a
38	55	68.8	328	12	Q82575	Q82575 influenza a
39	55	68.8	328	12	Q82576	Q82576 influenza a
40	55	68.8	328	12	Q82577	Q82577 influenza a
41	55	68.8	328	12	Q82582	Q82582 influenza a
42	55	68.8	328	12	Q82583	Q82583 influenza a
43	55	68.8	328	12	Q82586	Q82586 influenza a
44	55	68.8	328	12	Q82587	Q82587 influenza a
45	55	68.8	328	12	Q82589	Q82589 influenza a

ALIGNMENTS

```
RESULT   1
Q9W8Z1
ID   Q9W8Z1      PRELIMINARY;      PRT;   236 AA.
AC   Q9W8Z1;
DT   01-NOV-1999 (TrEMBLrel. 12, Created)
DT   01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   HEMAGGLUTININ (FRAGMENT).
OS   Influenza A virus.
OC   Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;
OC   Influenza virus A and B group; Influenza A viruses.
OX   NCBI_TaxID=11320;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=TOKYO1527, TOKYO1511;
RA   Mori S., Sekine H.;
RT   "HA1 domain of Influenza A (H3N2) virus.";
RL   Submitted (MAY-1998) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AB013807; BAA77285.1; -.
DR   EMBL; AB013806; BAA77284.1; -.
DR   InterPro; IPR001364; Hemagglutn.
DR   Pfam; PF00509; Hemagglutinin; 1.
DR   PRINTS; PR00329; HEMAGGLUTN12.
DR   ProDom; PD000225; Hemagglutn; 1.
FT   NON_TER       1      1
FT   NON_TER     236    236
SQ   SEQUENCE   236 AA;  26477 MW;  4566C8E7210FB558 CRC64;


  Query Match             68.8%;  Score 55;  DB 12;  Length 236;
  Best Local Similarity   100.0%;  Pred. No. 0.25;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 YPYDVPDYA 9
           |||||||||
Db      11 YPYDVPDYA 19


RESULT   2
Q9W8T9
ID   Q9W8T9      PRELIMINARY;      PRT;   236 AA.
AC   Q9W8T9;
```

```
DT   01-NOV-1999 (TrEMBLrel. 12, Created)
DT   01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   HEMAGGLUTININ (FRAGMENT).
OS   Influenza A virus.
OC   Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;
OC   Influenza virus A and B group; Influenza A viruses.
OX   NCBI_TaxID=11320;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=TOKYO1570, TOKYO1566, TOKYO1567, TOKYO1568, TOKYO1569;
RA   Mori S., Sekine H.;
RT   "HA1 domain of Influenza A (H3N2) virus.";
RL   Submitted (MAY-1998) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AB013813; BAA77291.1; -.
DR   EMBL; AB013809; BAA77287.1; -.
DR   EMBL; AB013810; BAA77288.1; -.
DR   EMBL; AB013811; BAA77289.1; -.
DR   EMBL; AB013812; BAA77290.1; -.
DR   InterPro; IPR001364; Hemagglutn.
DR   Pfam; PF00509; Hemagglutinin; 1.
DR   PRINTS; PR00329; HEMAGGLUTN12.
DR   ProDom; PD000225; Hemagglutn; 1.
FT   NON_TER       1      1
FT   NON_TER     236    236
SQ   SEQUENCE   236 AA;  26491 MW;  44CDD257210FB558 CRC64;

  Query Match             68.8%;  Score 55;  DB 12;  Length 236;
  Best Local Similarity   100.0%;  Pred. No. 0.25;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 YPYDVPDYA 9
          |||||||||
Db     11 YPYDVPDYA 19


RESULT   3
Q9W8J7
ID   Q9W8J7         PRELIMINARY;    PRT;   236 AA.
AC   Q9W8J7;
DT   01-NOV-1999 (TrEMBLrel. 12, Created)
DT   01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   HEMAGGLUTININ (FRAGMENT).
GN   HA.
OS   Influenza A virus.
OC   Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;
OC   Influenza virus A and B group; Influenza A viruses.
OX   NCBI_TaxID=11320;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=72, 70, AND 71;
RA   Nagashima M., Mori S., Sekine H.;
RT   "Influenza A virus gene for hemagglutinin.";
RL   Submitted (MAY-1998) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=TOKYO1539;
RA   Mori S., Sekine H.;
RT   "HA1 domain of Influenza A (H3N2) virus.";
RL   Submitted (MAY-1998) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AB014062; BAA77294.1; -.
DR   EMBL; AB013808; BAA77286.1; -.
DR   EMBL; AB014060; BAA77292.1; -.
DR   EMBL; AB014061; BAA77293.1; -.
DR   InterPro; IPR001364; Hemagglutn.
DR   Pfam; PF00509; Hemagglutinin; 1.
DR   PRINTS; PR00329; HEMAGGLUTN12.
DR   ProDom; PD000225; Hemagglutn; 1.
FT   NON_TER       1      1
FT   NON_TER     236    236
SQ   SEQUENCE   236 AA;  26493 MW;  457455682C8D28D7 CRC64;

  Query Match             68.8%;  Score 55;  DB 12;  Length 236;
  Best Local Similarity   100.0%;  Pred. No. 0.25;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 YPYDVPDYA 9
          |||||||||
Db     11 YPYDVPDYA 19


RESULT   4
Q9WI26
ID   Q9WI26      PRELIMINARY;      PRT;   286 AA.
AC   Q9WI26;
DT   01-NOV-1999 (TrEMBLrel. 12, Created)
DT   01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   H3 HEMAGGLUTININ (FRAGMENT).
GN   HA.
OS   influenza A virus (A/Taiwan/2034/96(H3N2)).
OC   Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;
OC   Influenza virus A and B group; Influenza A viruses; Influenza A virus.
OX   NCBI_TaxID=95241;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=A/TAIWAN/2034/96;
RA   Shih S.-R., Chen S.-H., Wu S.-L., Ho C.-C.;
RT   "Sequence Analysis of HA and NA Genes of Influenza Viruses Isolated in
RT   Taiwan.";
RL   Submitted (APR-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF139937; AAD34854.1; -.
DR   InterPro; IPR001364; Hemagglutn.
DR   Pfam; PF00509; Hemagglutinin; 1.
DR   PRINTS; PR00329; HEMAGGLUTN12.
DR   ProDom; PD000225; Hemagglutn; 1.
FT   NON_TER       1      1
FT   NON_TER     286    286
SQ   SEQUENCE   286 AA;  32075 MW;  7BA39C7632D33186 CRC64;

  Query Match             68.8%;  Score 55;  DB 12;  Length 286;
  Best Local Similarity   100.0%;  Pred. No. 0.31;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 YPYDVPDYA 9
          |||||||||
Db     57 YPYDVPDYA 65


RESULT   5
Q9WI25
ID   Q9WI25      PRELIMINARY;      PRT;   286 AA.
AC   Q9WI25;
DT   01-NOV-1999 (TrEMBLrel. 12, Created)
DT   01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   H3 HEMAGGLUTININ (FRAGMENT).
GN   HA.
OS   influenza A virus (A/Taiwan/2192/96(H3N2)).
OC   Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;
OC   Influenza virus A and B group; Influenza A viruses; Influenza A virus.
OX   NCBI_TaxID=95240;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=A/TAIWAN/2192/96;
RA   Shih S.-R., Chen S.-H., Wu S.-L., Ho C.-C.;
RT   "Sequence Analysis of HA and NA Genes of Influenza Viruses Isolated in
RT   Taiwan.";
RL   Submitted (APR-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF139936; AAD34853.1; -.
DR   InterPro; IPR001364; Hemagglutn.
DR   Pfam; PF00509; Hemagglutinin; 1.
```

```
DR   PRINTS; PR00329; HEMAGGLUTN12.
DR   ProDom; PD000225; Hemagglutn; 1.
FT   NON_TER       1      1
FT   NON_TER     286    286
SQ   SEQUENCE   286 AA;  32059 MW;  CB51B092AAD0D7D8 CRC64;


  Query Match             68.8%;  Score 55;  DB 12;  Length 286;
  Best Local Similarity   100.0%;  Pred. No. 0.31;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 YPYDVPDYA 9
          |||||||||
Db     57 YPYDVPDYA 65


RESULT   6
Q9WI24
ID   Q9WI24      PRELIMINARY;      PRT;   286 AA.
AC   Q9WI24;
DT   01-NOV-1999 (TrEMBLrel. 12, Created)
DT   01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   H3 HEMAGGLUTININ (FRAGMENT).
GN   HA.
OS   influenza A virus (A/Taiwan/2191/96(H3N2)).
OC   Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;
OC   Influenza virus A and B group; Influenza A viruses; Influenza A virus.
OX   NCBI_TaxID=95236;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=A/TAIWAN/2191/96;
RA   Shih S.-R., Chen S.-H., Wu S.-L., Ho C.-C.;
RT   "Sequence Analysis of HA and NA Genes of Influenza Viruses Isolated in
RT   Taiwan.";
RL   Submitted (APR-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF139932; AAD34849.1; -.
DR   InterPro; IPR001364; Hemagglutn.
DR   Pfam; PF00509; Hemagglutinin; 1.
DR   PRINTS; PR00329; HEMAGGLUTN12.
DR   ProDom; PD000225; Hemagglutn; 1.
FT   NON_TER       1      1
FT   NON_TER     286    286
SQ   SEQUENCE   286 AA;  31976 MW;  69AB2AE727D26078 CRC64;


  Query Match             68.8%;  Score 55;  DB 12;  Length 286;
  Best Local Similarity   100.0%;  Pred. No. 0.31;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 YPYDVPDYA 9
          |||||||||
Db     57 YPYDVPDYA 65


RESULT   7
O40774
ID   O40774      PRELIMINARY;      PRT;   325 AA.
AC   O40774;
DT   01-JAN-1998 (TrEMBLrel. 05, Created)
DT   01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   HEMAGGLUTININ GENE (FRAGMENT).
GN   HA.
OS   Influenza A virus.
OC   Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;
OC   Influenza virus A and B group; Influenza A viruses.
OX   NCBI_TaxID=11320;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=A/ILLINOIS/5/95(H3N2);
RA   Fitch W.M., Bush R.M., Bender C.A., Cox N.J.;
RL   Submitted (JUN-1997) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF008782; AAB63719.1; -.
DR   InterPro; IPR001364; Hemagglutn.
DR   Pfam; PF00509; Hemagglutinin; 1.
DR   ProDom; PD000225; Hemagglutn; 1.
FT   NON_TER       1      1
FT   NON_TER     325    325
SQ   SEQUENCE   325 AA;  36103 MW;  4E5B4648C2024765 CRC64;


  Query Match             68.8%;  Score 55;  DB 12;  Length 325;
  Best Local Similarity   100.0%;  Pred. No. 0.36;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 YPYDVPDYA 9
          |||||||||
Db     98 YPYDVPDYA 106


RESULT   8
Q9DL45
ID   Q9DL45      PRELIMINARY;      PRT;   325 AA.
AC   Q9DL45;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   HEMAGGLUTININ (FRAGMENT).
OS   Influenza A virus (A/Finland/583/98(H3N2)).
OC   Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;
OC   Influenza virus A and B group; Influenza A viruses; Influenza A virus.
OX   NCBI_TaxID=147149;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=A/FINLAND/583/98;
RA   Pyhala R., Haanpaa M., Kleemola M., Tervahauta R., Visakorpi R.,
RA   Kinnunen L.;
RT   "Acceptable protective efficacy of influenza vaccination in young
RT   military conscripts in circumstances of incomplete antigenic and
RT   genetic match.";
RL   Vaccine 0:0-0(2001).
DR   EMBL; AF311689; AAG47810.1; -.
DR   InterPro; IPR001364; Hemagglutn.
DR   Pfam; PF00509; Hemagglutinin; 1.
DR   PRINTS; PR00329; HEMAGGLUTN12.
DR   ProDom; PD000225; Hemagglutn; 1.
FT   NON_TER       1      1
FT   NON_TER     325    325
SQ   SEQUENCE   325 AA;  36029 MW;  5BECB8CC3E83B1C2 CRC64;


  Query Match             68.8%;  Score 55;  DB 12;  Length 325;
  Best Local Similarity   100.0%;  Pred. No. 0.36;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 YPYDVPDYA 9
          |||||||||
Db     98 YPYDVPDYA 106


RESULT   9
O40779
ID   O40779      PRELIMINARY;      PRT;   326 AA.
AC   O40779;
DT   01-JAN-1998 (TrEMBLrel. 05, Created)
DT   01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   HEMAGGLUTININ GENE (FRAGMENT).
GN   HA.
OS   Influenza A virus.
OC   Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;
OC   Influenza virus A and B group; Influenza A viruses.
OX   NCBI_TaxID=11320;
RN   [1]
RP   SEQUENCE FROM N.A.
```

```
RC   STRAIN=A/NEW_YORK/28/94(H3N2);
RA   Fitch W.M., Bush R.M., Bender C.A., Cox N.J.;
RL   Submitted (JUN-1997) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF008787; AAB63724.1; -.
DR   InterPro; IPR001364; Hemagglutn.
DR   Pfam; PF00509; Hemagglutinin; 1.
DR   ProDom; PD000225; Hemagglutn; 1.
FT   NON_TER       1      1
FT   NON_TER     326    326
SQ   SEQUENCE   326 AA;  36197 MW;  4B84BD4AADC44A77 CRC64;


  Query Match             68.8%;  Score 55;  DB 12;  Length 326;
  Best Local Similarity   100.0%;  Pred. No. 0.36;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 YPYDVPDYA 9
          |||||||||
Db     98 YPYDVPDYA 106


RESULT  10
O40780
ID   O40780      PRELIMINARY;      PRT;   326 AA.
AC   O40780;
DT   01-JAN-1998 (TrEMBLrel. 05, Created)
DT   01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   HEMAGGLUTININ GENE (FRAGMENT).
GN   HA.
OS   Influenza A virus.
OC   Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;
OC   Influenza virus A and B group; Influenza A viruses.
OX   NCBI_TaxID=11320;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=A/PENNSYLVANIA/7/94(H3N2);
RA   Fitch W.M., Bush R.M., Bender C.A., Cox N.J.;
RL   Submitted (JUN-1997) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF008788; AAB63725.1; -.
DR   InterPro; IPR001364; Hemagglutn.
DR   Pfam; PF00509; Hemagglutinin; 1.
DR   ProDom; PD000225; Hemagglutn; 1.
FT   NON_TER       1      1
FT   NON_TER     326    326
SQ   SEQUENCE   326 AA;  36349 MW;  6A11A4CBBFBD6CC9 CRC64;


  Query Match             68.8%;  Score 55;  DB 12;  Length 326;
  Best Local Similarity   100.0%;  Pred. No. 0.36;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 YPYDVPDYA 9
          |||||||||
Db     98 YPYDVPDYA 106


RESULT  11
O92324
ID   O92324      PRELIMINARY;      PRT;   327 AA.
AC   O92324;
DT   01-NOV-1998 (TrEMBLrel. 08, Created)
DT   01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   HEMAGGLUTININ (FRAGMENT).
GN   HA.
OS   Influenza A virus.
OC   Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;
OC   Influenza virus A and B group; Influenza A viruses.
OX   NCBI_TaxID=11320;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=A/SYDNEY/05/97-LIKE(H3N2);
RA   Osiowy C.K.;
RT   "Genetic characterization of A/Sydney/05/97-like(H3N2) Influenza virus
RT   isolates circulating in Canada during the 1997/98 Influenza season.";
RL   Submitted (AUG-1998) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF087700; AAC36729.1; -.
DR   InterPro; IPR001364; Hemagglutn.
DR   Pfam; PF00509; Hemagglutinin; 1.
DR   ProDom; PD000225; Hemagglutn; 1.
FT   NON_TER       1      1
FT   NON_TER     327    327
SQ   SEQUENCE   327 AA;  36389 MW;  909239953649069E CRC64;


  Query Match             68.8%;  Score 55;  DB 12;  Length 327;
  Best Local Similarity   100.0%;  Pred. No. 0.36;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 YPYDVPDYA 9
          |||||||||
Db     98 YPYDVPDYA 106


RESULT  12
O92325
ID   O92325      PRELIMINARY;      PRT;   327 AA.
AC   O92325;
DT   01-NOV-1998 (TrEMBLrel. 08, Created)
DT   01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   HEMAGGLUTININ (FRAGMENT).
GN   HA.
OS   Influenza A virus.
OC   Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;
OC   Influenza virus A and B group; Influenza A viruses.
OX   NCBI_TaxID=11320;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=A/SYDNEY/05/97-LIKE(H3N2);
RA   Osiowy C.K.;
RT   "Genetic characterization of A/Sydney/05/97-like(H3N2) Influenza virus
RT   isolates circulating in Canada during the 1997/98 Influenza season.";
RL   Submitted (AUG-1998) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF087703; AAC36732.1; -.
DR   InterPro; IPR001364; Hemagglutn.
DR   Pfam; PF00509; Hemagglutinin; 1.
DR   ProDom; PD000225; Hemagglutn; 1.
FT   NON_TER       1      1
FT   NON_TER     327    327
SQ   SEQUENCE   327 AA;  36457 MW;  D741E3C39D0C4D6A CRC64;


  Query Match             68.8%;  Score 55;  DB 12;  Length 327;
  Best Local Similarity   100.0%;  Pred. No. 0.36;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 YPYDVPDYA 9
          |||||||||
Db     98 YPYDVPDYA 106


RESULT  13
O92327
ID   O92327      PRELIMINARY;      PRT;   327 AA.
AC   O92327;
DT   01-NOV-1998 (TrEMBLrel. 08, Created)
DT   01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   HEMAGGLUTININ (FRAGMENT).
GN   HA.
OS   Influenza A virus.
OC   Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;
OC   Influenza virus A and B group; Influenza A viruses.
OX   NCBI_TaxID=11320;
```

```
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=A/SYDNEY/05/97(H3N2);
RA   Osiowy C.K.;
RT   "Genetic characterization of A/Sydney/05/97-like(H3N2) Influenza virus
RT   isolates circulating in Canada during the 1997/98 Influenza season.";
RL   Submitted (AUG-1998) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF087709; AAC36738.1; -.
DR   InterPro; IPR001364; Hemagglutn.
DR   Pfam; PF00509; Hemagglutinin; 1.
DR   ProDom; PD000225; Hemagglutn; 1.
FT   NON_TER       1      1
FT   NON_TER     327    327
SQ   SEQUENCE   327 AA;  36293 MW;  4A72007829B091F5 CRC64;


  Query Match             68.8%;  Score 55;  DB 12;  Length 327;
  Best Local Similarity   100.0%;  Pred. No. 0.36;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 YPYDVPDYA 9
          |||||||||
Db     98 YPYDVPDYA 106


RESULT  14
Q9YT90
ID   Q9YT90      PRELIMINARY;      PRT;   327 AA.
AC   Q9YT90;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   HEMAGGLUTININ (FRAGMENT).
GN   HA.
OS   influenza A virus (A/Sydney/05/97-like(H3N2)).
OC   Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;
OC   Influenza virus A and B group; Influenza A viruses; Influenza A virus.
OX   NCBI_TaxID=82372;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=A/SYDNEY/05/97-LIKE(H3N2);
RA   Osiowy C.K.;
RT   "Genetic characterization of A/Sydney/05/97-like(H3N2) Influenza virus
RT   isolates circulating in Canada during the 1997/98 Influenza season.";
RL   Submitted (AUG-1998) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF087701; AAC36730.1; -.
DR   InterPro; IPR001364; Hemagglutn.
DR   Pfam; PF00509; Hemagglutinin; 1.
DR   PRINTS; PR00329; HEMAGGLUTN12.
DR   ProDom; PD000225; Hemagglutn; 1.
FT   NON_TER       1      1
FT   NON_TER     327    327
SQ   SEQUENCE   327 AA;  36373 MW;  ADBD67D856EBCC96 CRC64;


  Query Match             68.8%;  Score 55;  DB 12;  Length 327;
  Best Local Similarity   100.0%;  Pred. No. 0.36;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 YPYDVPDYA 9
          |||||||||
Db     98 YPYDVPDYA 106


RESULT  15
Q9YT89
ID   Q9YT89      PRELIMINARY;      PRT;   327 AA.
AC   Q9YT89;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   HEMAGGLUTININ (FRAGMENT).
GN   HA.
OS   influenza A virus (A/Sydney/05/97-like(H3N2)).
OC   Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;
OC   Influenza virus A and B group; Influenza A viruses; Influenza A virus.
OX   NCBI_TaxID=82372;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=A/SYDNEY/05/97-LIKE(H3N2);
RA   Osiowy C.K.;
RT   "Genetic characterization of A/Sydney/05/97-like(H3N2) Influenza virus
RT   isolates circulating in Canada during the 1997/98 Influenza season.";
RL   Submitted (AUG-1998) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF087702; AAC36731.1; -.
DR   InterPro; IPR001364; Hemagglutn.
DR   Pfam; PF00509; Hemagglutinin; 1.
DR   PRINTS; PR00329; HEMAGGLUTN12.
DR   ProDom; PD000225; Hemagglutn; 1.
FT   NON_TER       1      1
FT   NON_TER     327    327
SQ   SEQUENCE   327 AA;  36446 MW;  9E9FA8433CFC1B01 CRC64;


  Query Match             68.8%;  Score 55;  DB 12;  Length 327;
  Best Local Similarity   100.0%;  Pred. No. 0.36;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 YPYDVPDYA 9
          |||||||||
Db     98 YPYDVPDYA 106


Search completed: April 11, 2002, 10:22:57
Job time: 427 sec
```

THIS PAGE BLANK (USPTO)

GenCore version 4.5

OM protein - protein search, using sw model

Run on: April 11, 2002, 10:15:43 ; Search time 53.05 Seconds
(without alignments)
18.152 Million cell updates/sec

Title: US-09-284-787-3
Perfect score: 75
Sequence: 1 SGSGYPYDVPDYA 13

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 522463 seqs, 74073290 residues

Total number of hits satisfying chosen parameters: 522463

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_1101:*
1: /SIDS8/gcgdata/geneseq/geneseqp/AA1980.DAT:*
2: /SIDS8/gcgdata/geneseq/geneseqp/AA1981.DAT:*
3: /SIDS8/gcgdata/geneseq/geneseqp/AA1982.DAT:*
4: /SIDS8/gcgdata/geneseq/geneseqp/AA1983.DAT:*
5: /SIDS8/gcgdata/geneseq/geneseqp/AA1984.DAT:*
6: /SIDS8/gcgdata/geneseq/geneseqp/AA1985.DAT:*
7: /SIDS8/gcgdata/geneseq/geneseqp/AA1986.DAT:*
8: /SIDS8/gcgdata/geneseq/geneseqp/AA1987.DAT:*
9: /SIDS8/gcgdata/geneseq/geneseqp/AA1988.DAT:*
10: /SIDS8/gcgdata/geneseq/geneseqp/AA1989.DAT:*
11: /SIDS8/gcgdata/geneseq/geneseqp/AA1990.DAT:*
12: /SIDS8/gcgdata/geneseq/geneseqp/AA1991.DAT:*
13: /SIDS8/gcgdata/geneseq/geneseqp/AA1992.DAT:*
14: /SIDS8/gcgdata/geneseq/geneseqp/AA1993.DAT:*
15: /SIDS8/gcgdata/geneseq/geneseqp/AA1994.DAT:*
16: /SIDS8/gcgdata/geneseq/geneseqp/AA1995.DAT:*
17: /SIDS8/gcgdata/geneseq/geneseqp/AA1996.DAT:*
18: /SIDS8/gcgdata/geneseq/geneseqp/AA1997.DAT:*
19: /SIDS8/gcgdata/geneseq/geneseqp/AA1998.DAT:*
20: /SIDS8/gcgdata/geneseq/geneseqp/AA1999.DAT:*
21: /SIDS8/gcgdata/geneseq/geneseqp/AA2000.DAT:*
22: /SIDS8/gcgdata/geneseq/geneseqp/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	75	100.0	13	19	AAW59452	Human influenza vi
2	71	94.7	34	21	AAB07730	Peptide encoded by
3	67	89.3	17	18	AAW41032	Flu-epitope scramb
4	67	89.3	17	18	AAW29116	Flu epitope tagged
5	67	89.3	17	22	AAB45969	Transdominant effe
6	67	89.3	19	21	AAY90725	Influenza haemaggl
7	67	89.3	38	20	AAY49751	Compact structure
8	65	86.7	16	21	AAB03963	pTAT-HA linker pol
9	65	86.7	16	21	AAB29443	pTAT-HA linker-enc
10	64.5	86.0	165	22	AAG63606	Amino acid sequenc
11	62	82.7	481	18	AAW27158	Saccharomyces cere
12	61	81.3	10	20	AAW73372	Haemagglutinin epi
13	61	81.3	11	20	AAW73371	Haemagglutinin epi
14	61	81.3	11	22	AAB73290	Haemagglutinin, HA
15	61	81.3	12	20	AAW73370	Haemagglutinin epi
16	61	81.3	12	22	AAB49724	Influenza haemaglu
17	61	81.3	80	21	AAY97533	Truncated Human p2
18	61	81.3	80	21	AAY96048	Truncated cyclin d
19	61	81.3	80	21	AAY96075	Human cyclin depen
20	61	81.3	177	21	AAY97530	Truncated Human p2
21	61	81.3	177	21	AAY96045	Human truncated cy
22	61	81.3	177	21	AAY96072	Human cyclin depen
23	61	81.3	224	22	AAB99371	Human interleukin
24	61	81.3	224	22	AAB99372	Human interleukin
25	61	81.3	224	22	AAB99373	Human interleukin
26	61	81.3	224	22	AAB99375	Human interleukin
27	61	81.3	224	22	AAB99377	Human interleukin
28	61	81.3	224	22	AAB99379	Human interleukin
29	61	81.3	224	22	AAB99392	Human interleukin
30	61	81.3	224	22	AAB99393	Human interleukin
31	61	81.3	224	22	AAB99395	Human interleukin
32	61	81.3	224	22	AAB99396	Human interleukin
33	61	81.3	224	22	AAB99398	Human interleukin
34	61	81.3	224	22	AAB99399	Human interleukin
35	61	81.3	224	22	AAB75004	Anti-IL8 monoclona
36	61	81.3	224	22	AAB75005	Anti-IL8 monoclona
37	61	81.3	224	22	AAB75006	Anti-IL8 monoclona
38	61	81.3	224	22	AAB75008	Anti-IL8 monoclona
39	61	81.3	224	22	AAB75010	Anti-IL8 monoclona
40	61	81.3	224	22	AAB75012	Anti-IL8 monoclona
41	61	81.3	224	22	AAB75025	Anti-IL8 monoclona
42	61	81.3	224	22	AAB75026	Anti-IL8 monoclona
43	61	81.3	224	22	AAB75028	Anti-IL8 monoclona
44	61	81.3	224	22	AAB75029	Anti-IL8 monoclona
45	61	81.3	224	22	AAB75031	Anti-IL8 monoclona

ALIGNMENTS

RESULT 1
AAW59452
ID AAW59452 standard; peptide; 13 AA.
XX
AC AAW59452;
XX
DT 28-AUG-1998 (first entry)
XX
DE Human influenza virus haemagglutinin epitope peptide fragment #3.
XX
KW Haemagglutinin; epitope; monoclonal antibody; detection; isolation;
KW fusion protein.
XX
OS Human parainfluenza virus.
XX
FH Key Location/Qualifiers
FT Modified-site 1
FT /note= "N-terminal biotinoyl-epsilon-Aca group"
FT Modified-site 13
FT /note= "C-terminal amide group"
XX
PN DE19643314-A1.
XX
PD 23-APR-1998.
XX
PF 21-OCT-1996; 96DE-1043314.
XX
PR 21-OCT-1996; 96DE-1043314.
XX
PA (BOEF) BOEHRINGER MANNHEIM GMBH.
XX
PI Emrich T, Grol M, Hinzpeter M;
XX
DR WPI; 1998-241603/22.

```
XX
PT   Monoclonal antibodies to influenza virus haemagglutinin - useful for
PT   detecting and isolating haemagglutinin proteins
XX
PS   Claim 7; Page 2; 6pp; German.
XX
CC   This peptide fragment represents a human influenza virus haemagglutinin
CC   epitope which is used to make monoclonal antibodies with an affinity
CC   constant of 108 M-1. Such antibodies are used for detecting and isolating
CC   native human influenza virus haemagglutinin (HA), modified HA or HA
CC   fusion proteins.
XX
SQ   Sequence   13 AA;


  Query Match             100.0%;  Score 75;  DB 19;  Length 13;
  Best Local Similarity   100.0%;  Pred. No. 1.6e-05;
  Matches   13;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 SGSGYPYDVPDYA 13
          |||||||||||||
Db      1 sgsgypydvpdya 13


RESULT   2
AAB07730
ID   AAB07730 standard; Peptide; 34 AA.
XX
AC   AAB07730;
XX
DT   07-NOV-2000  (first entry)
XX
DE   Peptide encoded by amplification product encoding a FLAG-HA peptide.
XX
KW   Genetic element; optical property; gene product; microcapsule.
XX
OS   Synthetic.
XX
PN   WO200040712-A1.
XX
PD   13-JUL-2000.
XX
PF   06-JAN-2000; 2000WO-GB00030.
XX
PR   07-JAN-1999;   99GB-0000298.
XX
PA   (MEDI-) MEDICAL RES COUNCIL.
XX
PI   Griffiths A,  Tawfik D;
XX
DR   WPI; 2000-465978/40.
DR   N-PSDB; AAA59266.
XX
PT   Isolation of genetic elements encoding a gene product with a desired
PT   activity is carried out in microcapsules where changes in optical
PT   activity are detected -
XX
PS   Example 12; Page 79-80; 128pp; English.
XX
CC   The specification describes a method for isolating one or more genetic
CC   elements encoding a gene product with a desired activity. Expression of
CC   the genetic element results, directly or indirectly in the modification
CC   of an optical property of the encoded polypeptide. The method comprises
CC   compartmentalizing the genetic elements into microcapsules, expressing
CC   them to produce their respective gene products within the microcapsules,
CC   and sorting genetic elements which produce the gene product(s) with the
CC   desired activity. The method is used for isolating genetic elements
CC   which encode a gene product with a desired activity. This can be a
CC   catalytic, regulatory or binding activity. The present sequence
CC   represents a fusion of a flag epitope and influenza haemagglutinin. It
CC   was used in the course of the invention.
XX
SQ   Sequence   34 AA;


  Query Match             94.7%;  Score 71;  DB 21;  Length 34;
  Best Local Similarity   92.3%;  Pred. No. 0.00018;
  Matches   12;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      1 SGSGYPYDVPDYA 13
          || ||||||||||
Db     18 sgggypydvpdya 30


RESULT   3
AAW41032
ID   AAW41032 standard; Protein; 17 AA.
XX
AC   AAW41032;
XX
DT   12-MAY-1998  (first entry)
XX
DE   Flu-epitope scrambled peptide.
XX
KW   Transdominant bioactive agent; phenotype alteration; signalling pathway;
KW   antitumour agent; cardiovascular disease; angiogenesis; atherosclerosis;
KW   obesity; neurodegeneration; bone disease; infection; allergy; therapy;
KW   flu-epitope peptide.
XX
OS   Synthetic.
OS   Influenza virus.
XX
PN   WO9727213-A1.
XX
PD   31-JUL-1997.
XX
PF   23-JAN-1997;   97WO-US01048.
XX
PR   23-JAN-1996;   96US-0589911.
PR   23-JAN-1996;   96US-0589109.
XX
PA   (STRD ) UNIV LELAND STANFORD JUNIOR.
XX
PI   Noaln GP,  Rothenberg SM;
XX
DR   WPI; 1997-393612/36.
XX
PT   Screening for trans-dominant active agents able to alter cell
PT   phenotype - useful to identify potential drugs with e. g.
PT   anti-tumour activity
XX
PS   Example 1; Page 53; 91pp; English.
XX
CC   This sequence represents a scrambled flu-epitope peptide, and can be
CC   used in the method of the invention. The method is for screening for
CC   transdominant bioactive agents (A) able to alter the phenotype of a cell
CC   comprises: (a) introducing a library of different randomised nucleic
CC   acids (I) into cells; and (b) screening the cells for altered phenotype
CC   due to presence of (A). (A) are potential pharmaceuticals, from their
CC   effect on signalling pathways. A particular application is to identify
CC   peptides, encoded by (I), that have antitumour activity or are able to
CC   increase sensitivity to, or reduce toxicity of, known antitumour agents.
CC   More generally the method can identify agents for treatment of many other
CC   conditions, e.g. cardiovascular diseases, angiogenesis, atherosclerosis,
CC   obesity, neurodegeneration, bone disease, viral and other infections,
CC   allergy etc.
XX
SQ   Sequence   17 AA;


  Query Match             89.3%;  Score 67;  DB 18;  Length 17;
  Best Local Similarity   91.7%;  Pred. No. 0.00038;
  Matches   11;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      2 GSGYPYDVPDYA 13
          | ||||||||||
```

```
Db      2 gggypydvpdya 13


RESULT   4
AAW29116
ID   AAW29116 standard; Peptide; 17 AA.
XX
AC   AAW29116;
XX
DT   04-FEB-1998  (first entry)
XX
DE   Flu epitope tagged scrambled peptide.
XX
KW   Transdominant intracellular effector peptide; RNA; screening;
KW   altered phenotype; retrovirus; library; signal transduction;
KW   antitumour; pBABE; influenza virus; flu; epitope.
XX
OS   Synthetic.
XX
PN   WO9727212-A1.
XX
PD   31-JUL-1997.
XX
PF   23-JAN-1997;   97WO-US01019.
XX
PR   23-JAN-1996;   96US-0589911.
PR   23-JAN-1996;   96US-0589109.
XX
PA   (RIGE-) RIGEL PHARM INC.
XX
PI   Noaln GP;
XX
DR   WPI; 1997-393611/36.
XX
PT   Screening for transdominant intracellular active agents able to
PT   alter cell phenotype - useful for examining for changed phenotype,
PT   particularly to identify potential drugs with e. g. antitumour
PT   activity
XX
PS   Example 1; Page 53; 91pp; English.
XX
CC   This peptide comprises a Flu epitope tagged scrambled peptide.
CC   Murine 10T1/2 Clone 8 cells were infected with pBabe puro
CC   retrovirus constructs containing a sequence coding for the Flu
CC   scrambled epitope or an inhibitor peptide (see AAW29117) identical
CC   to the C2 region of protein kinase C (PKA) isozymes. Control clone
CC   8 cells showed predominantly cytoplasmic and perinuclear staining,
CC   while PMA induced cells showed translocation to the nucleus. Cells
CC   infected with constructs coding for the scrambled peptide showed
CC   similar staining. Those infected with constructs coding for
CC   the C2 region showed predominantly cytoplasmic and perinuclear
CC   staining in both control and PMA induced cells. Novel outcomes can
CC   occur upon expression of peptides in cells. A claimed method of
CC   screening for transdominant intracellular bioactive agents (A) able
CC   to alter the phenotype of a cell comprises: (a) introducing a
CC   library of different randomised nucleic acids (I) into cells; and
CC   (b) screening the cells for altered phenotype due to presence of
CC   (A). Also new are: (1) a library of retroviruses containing
CC   different randomised (I); and (2) a library of mammalian cells
CC   containing the library of (1), preferably integrated into the
CC   genome.
XX
SQ   Sequence   17 AA;


  Query Match             89.3%;  Score 67;  DB 18;  Length 17;
  Best Local Similarity   91.7%;  Pred. No. 0.00038;
  Matches   11;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      2 GSGYPYDVPDYA 13
          | ||||||||||
Db      2 gggypydvpdya 13


RESULT   5
AAB45969
ID   AAB45969 standard; Peptide; 17 AA.
XX
AC   AAB45969;
XX
DT   02-APR-2001  (first entry)
XX
DE   Transdominant effector peptide associated screening peptide #49.
XX
KW   Intracellular transdominant bioactive agent; screening; cell phenotype;
KW   effector peptide.
XX
OS   Unidentified.
XX
PN   US6153380-A.
XX
PD   28-NOV-2000.
XX
PF   23-JAN-1997;   97US-0789333.
XX
PR   23-JAN-1996;   96US-0589108.
PR   23-JAN-1996;   96US-0589911.
XX
PA   (STRD ) UNIV LELAND STANFORD JUNIOR.
PA   (RIGE-) RIGEL PHARM INC.
XX
PI   Rothenberg SM,  Nolan GP;
XX
DR   WPI; 2001-060084/07.
XX
PT   Methods for screening intracellular transdominant effector peptides and
PT   RNA molecules comprise delivering random oligonucleotides to cells,
PT   which are then screened for an altered phenotype  -
XX
PS   Example 1; Column 79-80; 57pp; English.
XX
CC   This invention describes novel in vitro screening methods (I) for a
CC   transdominant intracellular bioactive agent capable of altering the
CC   phenotype of a cell. (I) comprises: (a) introducing a molecular library
CC   of randomized candidate nucleic acids into several cells; and (b)
CC   screening the cells for a cell exhibiting an altered phenotype, where the
CC   altered phenotype is due to the presence of a transdominant bioactive
CC   agent. The methods are particularly useful for screening intracellular
CC   transdominant effector peptides and RNA molecules selected inside living
CC   cells from randomized pools. (I) is also useful for introducing random
CC   libraries into cells to screen for bioactive compounds. The methods allow
CC   rapid and highly efficient screening of large numbers of random
CC   oligonucleotides and their corresponding expression products in a single
CC   step. In addition, the methods allow screening in the absence of
CC   significant prior characterization of the cellular defect.
XX
SQ   Sequence   17 AA;


  Query Match             89.3%;  Score 67;  DB 22;  Length 17;
  Best Local Similarity   91.7%;  Pred. No. 0.00038;
  Matches   11;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      2 GSGYPYDVPDYA 13
          | ||||||||||
Db      2 gggypydvpdya 13


RESULT   6
AAY90725
ID   AAY90725 standard; peptide; 19 AA.
XX
AC   AAY90725;
XX
DT   16-AUG-2000  (first entry)
XX
```

DE Influenza haemagglutinin epitope tag containing 19mer peptide.
XX
KW Fusion; scaffold protein; peptide library; green fluorescent protein;
KW GFP; self-binding; detection; cellular expression; screening; catabolism.
XX
OS Influenza virus.
XX
PN WO200020574-A2.
XX
PD 13-APR-2000.
XX
PF 08-OCT-1999; 99WO-US23715.
XX
PR 08-OCT-1998; 98US-0169015.
XX
PA (RIGE-) RIGEL PHARM INC.
XX
PI Anderson D, Bogenberger JM, Peelle BR;
XX
DR WPI; 2000-303765/26.
XX
PT Fusions of scaffold proteins with random peptide libraries for improved
PT library screening -
XX
PS Example 2; Page 80; 89pp; English.
XX
CC The present invention describes a library (X) of fusion nucleic acids,
CC each comprising a nucleic acid encoding a random peptide, a nucleic
CC acid encoding a scaffold protein and a nucleic acid encoding a fusion
CC partner (the nucleic acid encoding the random peptide is inserted
CC internally into the scaffold protein nucleic acid). Also described are:
CC (1) a library of retroviral vectors comprising (X); (2) a library of
CC host cells comprising (X); and (3) a method (METH1) of screening for
CC bioactive peptides conferring a particular phenotype, comprising
CC providing cells containing (X). The constructs may be used in the
CC screening of peptide libraries. The use of the scaffold proteins
CC (especially green fluorescent protein (GFP)) in the fusion constructs
CC increases cellular expression levels, increases conformational stability
CC relative to linear peptides, decreases cellular catabolism, and
CC increases the steady state concentrations of the random peptides and
CC random peptide library members expressed in cells for the purposes of
CC detecting them. The present sequence represents a peptide which
CC comprises the influenza haemagglutinin (HA) epitope tag embedded with
CC glycines, which is used in an example from the present invention.
XX
SQ Sequence 19 AA;

Query Match 89.3%; Score 67; DB 21; Length 19;
Best Local Similarity 91.7%; Pred. No. 0.00042;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 GSGYPYDVPDYA 13
| ||||||||||
Db 2 gggypydvpdya 13

RESULT 7
AAY49751
ID AAY49751 standard; peptide; 38 AA.
XX
AC AAY49751;
XX
DT 19-JAN-2000 (first entry)
XX
DE Compact structure forming exemplification peptide #44.
XX
KW Compact structure forming peptide; dimerisation; stability; scaffold;
KW library screening; drug screening; gene therapy.
XX
OS Synthetic.
XX
PN WO9951625-A2.
XX
PD 14-OCT-1999.
XX
PF 02-APR-1999; 99WO-US07374.
XX
PR 02-APR-1998; 98US-0080444.
XX
PA (RIGE-) RIGEL PHARM INC.
XX
PI Anderson D;
XX
DR WPI; 1999-620191/53.
XX
PT Novel dimerization peptides which self-associate are used with other
PT proteins to effect the formation of compact structures -
XX
PS Disclosure; Page 13; 75pp; English.
XX
CC The present invention describes peptides which have a moderate or high
CC affinity for each other, when added as extensions to both the N-terminus
CC and C-terminus of a protein, can be used to help fold the protein into
CC a compact structure. This compact structure is more stable to proteases.
CC A composition (A) comprises at least a first dimerisation peptide (I)
CC comprising the sequence (S1) that is no more than 8 amino acids long,
CC where the composition optionally comprises a second dimerisation peptide
CC (II): NH2-X1-X2-X3-X4-X5-COOH (S1) where X1 to X4 = Ala, Val, Ile, Leu,
CC Trp, Phe, Met, or Tyr, and X5 = Lys, Arg, Asp or Glu;
CC NH2-X1-X2-X3-X4-X5-COOH (II). The compositions of the invention are
CC displayed intracellularly or extracellularly and are useful to identity
CC binding proteins and molecules, and to modulate intracellular signalling
CC pathways. A library of constrained proteins may be evaluated in vivo for
CC its bioactive potential. The invention can be used to access molecules
CC or targets within living cells, and then provide for the isolation of
CC the constrained protein which has a phenotypic effect on the living
CC cells. The methods are also useful to identify in vitro binding partners
CC of the constrained protein. The compositions of the invention are useful
CC as a scaffold for gene therapy and for potential use as a therapeutic
CC in physiological fluids. The present sequence is used in the
CC exemplification of the present invention.
XX
SQ Sequence 38 AA;

Query Match 89.3%; Score 67; DB 20; Length 38;
Best Local Similarity 91.7%; Pred. No. 0.00087;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 GSGYPYDVPDYA 13
| ||||||||||
Db 11 gggypydvpdya 22

RESULT 8
AAB03963
ID AAB03963 standard; Peptide; 16 AA.
XX
AC AAB03963;
XX
DT 26-FEB-2001 (first entry)
XX
DE pTAT-HA linker polypeptide.
XX
KW Chimeric protein; fusion protein; FLICE like inhibitor protein;
KW FLIP; Fas; TNF; apoptosis; caspase-8; ligand; T cell; thymocyte;
KW tumour specific antigen; immune response; therapy; prophylaxis;
KW diagnosis; HIV; human immunodeficiency syndrome; AIDS;
KW acquired immune deficiency syndrome.
XX
OS Human immunodeficiency virus.
XX
PN WO200059935-A1.
XX
PD 12-OCT-2000.

XX
PF 05-APR-2000; 2000WO-US09002.
XX
PR 05-APR-1999; 99US-0127867.
PR 06-APR-1999; 99US-0128021.
XX
PA (MAYO-) MAYO FOUND MEDICAL EDUCATION & RES.
PA (PAYA/) PAYA C.
PA (ALGE/) ALGECIRAS-SCHMINICH A.
XX
PI Paya C, Algeciras-schminich A;
XX
DR WPI; 2000-664988/64.
DR N-PSDB; AAA54300.
XX
PT Fusion polypeptide useful for inhibiting ligand-induced apoptosis,
PT comprises portion of anti-apoptotic polypeptide linked to a transport
PT group
XX
PS Disclosure; Fig 9; 89pp; English.
XX
CC A chimeric group or fusion peptide which comprises a portion of an
CC anti-apoptotic polypeptide which inhibits apoptosis of lymphocytes in
CC combination with a transport group is described. The transport group
CC is capable of transporting the chimeric group or fusion peptide
CC across the cell membrane. The anti-apoptotic polypeptide is FLICE-like
CC inhibitor protein (FLIP) which inhibits Fas and TNF mediated apoptosis
CC by inhibiting binding of Caspase-8 to the Fas receptor complex, thus
CC shutting off the downstream Fas signalling pathway. The chimeric group
CC and fusion peptide are useful for inhibiting ligand-induced apoptosis
CC by bringing them into contact with T cells. The chimeric group is
CC useful for expanding T cells in vitro e.g. T cells specific for
CC particular antigens such as tumour-specific antigen, for enhancing
CC immune response and to inhibit the apoptosis of chronically activated
CC T cells e.g. activated CD4^+ T cells in HIV infected patients. The
CC chimeric group is also useful for therapeutic, prophylactic or
CC diagnosis of intracellular delivery of small molecules and
CC macromolecules such as anti-apoptotic polypeptides and nucleic
CC acids encoding such polypeptides. Two primers (AAA54297, AAA54298) were
CC used to amplify the FLIP cDNA for subcloning into the XhoI-NcoI
CC site of the pHA-TAT vector which contains the N-terminal protein
CC transduction domain from the human immunodeficiency virus tat
CC protein. Tat is a preferred transport moiety.
XX
SQ Sequence 16 AA;

Query Match 86.7%; Score 65; DB 21; Length 16;
Best Local Similarity 100.0%; Pred. No. 0.00073;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 SGYPYDVPDYA 13
||||||||||||
Db 2 sgypydvpdya 12

RESULT 9
AAB29443
ID AAB29443 standard; peptide; 16 AA.
XX
AC AAB29443;
XX
DT 09-FEB-2001 (first entry)
XX
DE pTAT-HA linker-encoded peptide.
XX
KW Protein transduction domain; fusion molecule; therapeutic agent;
KW drug targetting; drug discovery; cell transduction; bioavailability;
KW vaccine; nervous system disorder; Alzheimer's disease;
KW Parkinson's disease; Huntington's disease; pre-senile dementia; epilepsy;
KW seizure; compulsive behaviour; meningitis; encephalitis; ischaemia;
KW spongiform encephalopathy; dyslexia; age-related memory loss;
KW Lou Gehring's disease; viral infection; HIV; bacterial infection.
XX
OS Human immunodeficiency virus.
OS Synthetic.
XX
PN WO200062067-A1.
XX
PD 19-OCT-2000.
XX
PF 28-FEB-2000; 2000WO-US05097.
XX
PR 28-FEB-1999; 99US-0122757.
PR 29-AUG-1999; 99US-0151291.
XX
PA (UNIW) UNIV WASHINGTON.
XX
PI Dowdy SF;
XX
DR WPI; 2000-647439/62.
DR N-PSDB; AAC63859.
XX
PT Fusion molecules comprising protein transduction domains and
PT therapeutic agents, useful for treating e.g. Alzheimer's and
PT Parkinson's diseases, dementia and epilepsy -
XX
PS Example 1; Fig 2; 191pp; English.
XX
CC The invention relates to a novel fusion molecule comprising at least
CC one protein transduction domain (PTD) and at least one linked molecule,
CC where the linked molecule has therapeutic or prophylactic activity
CC against a medical condition. The invention also relates to methods of
CC drug discovery in which the test compound is linked to a suitable
CC transducing protein and introduced to a cell; a method of killing
CC resistant microorganisms using a suitable fusion molecule; a mammal
CC comprising a covalently linked fusion molecule; and a mammal adapted for
CC experimental use in which at least one transduction molecule has been
CC transduced into essentially all the cells of the mammal. The fusion
CC molecule is used to deliver a therapeutic agent to a mammal, especially
CC a human. The linked molecule may be a vaccine, an anti-infective drug,
CC a cardiovascular drug, an antitumour drug, an analgesic, an
CC antiinflammatory, a diagnostic marker or a drug for the treatment or
CC prevention of a central or peripheral nervous system disorder. The
CC central nervous system (CNS) disorder is especially Alzheimer's disease,
CC Parkinson's disease, Huntington's disease, and also includes pre-senile
CC dementia, epilepsy and seizures, compulsive behaviour, meningitis
CC (including viral and bacterial meningitis), encephalitis, ischaemia,
CC scrapie (or related spongiform encephalopathies), dyslexia, age-related
CC memory loss or Lou Gehring's disease. Fusion molecules can also be
CC used to kill virally infected cells, especially those infected with HIV.
CC The vaccines are used to treat or prevent bacterial or viral infections.
CC The methods are a highly effective means for transducing a molecule
CC into an entire mammal or into specific cells, tissues, organs and
CC systems within it. They also overcome bioavailability problems that
CC are associated with many therapeutic agents (e.g., large molecular size,
CC hydrophobicity, hydrophilicity, biological resistance), by providing
CC efficient transduction of the target cell. The present sequence
CC represents a protein transduction domain used in the invention.
XX
SQ Sequence 16 AA;

Query Match 86.7%; Score 65; DB 21; Length 16;
Best Local Similarity 100.0%; Pred. No. 0.00073;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 SGYPYDVPDYA 13
||||||||||||
Db 2 sgypydvpdya 12

RESULT 10
AAG63606
ID AAG63606 standard; Protein; 165 AA.
XX

AC AAG63606;
XX
DT 15-OCT-2001 (first entry)
XX
DE Amino acid sequence of C-terminal fusion protein of pIX protein.
XX
KW pIX protein; chimeric protein; gene transfer vector.
XX
OS Synthetic.
OS Mastadenovirus.
XX
PN WO200158940-A2.
XX
PD 16-AUG-2001.
XX
PF 09-FEB-2001; 2001WO-US04233.
XX
PR 09-FEB-2000; 2000US-0181163.
XX
PA (GENV-) GENVEC INC.
XX
PI Roelvink PW, Kovesdi I, Wickham TJ;
XX
DR WPI; 2001-497066/54.
DR N-PSDB; AAH74659.
XX
PT Chimeric pIX protein useful in an adenovirus gene transfer vector for
PT infecting cells comprises at least one adenoviral pIX domain and a
PT non-native amino acid -
XX
PS Example 1; Page 27-28; 28pp; English.
XX
CC The present sequence represents a chimeric fusion protein of the
CC invention. The specification describes a chimeric pIX protein having
CC at least one adenoviral pIX domain and a non-native amino acid sequence.
CC The non-native sequence is a ligand that binds to a substrate present on
CC surface cells. The chimeric proteins are used for producing adenovirus
CC gene transfer vectors.
XX
SQ Sequence 165 AA;

Query Match 86.0%; Score 64.5; DB 22; Length 165;
Best Local Similarity 92.9%; Pred. No. 0.0098;
Matches 13; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

Qy 1 SGSG-YPYDVPDYA 13
|||| |||||||||
Db 150 sgsgsypydvpdya 163

RESULT 11
AAW27158
ID AAW27158 standard; Protein; 481 AA.
XX
AC AAW27158;
XX
DT 20-APR-1998 (first entry)
XX
DE Saccharomyces cerevisiae alpha mating factor receptor Ste2.
XX
KW Alpha mating factor receptor; Ste2; mutant; biosensor; diagnosis;
KW G-protein coupled receptor; GPCR; environmental monitoring;
KW on-line process monitoring.
XX
OS Saccharomyces cerevisiae.
XX
PN WO9735985-A1.
XX
PD 02-OCT-1997.
XX
PF 19-MAR-1997; 97WO-GB00746.
XX
PR 22-MAR-1996; 96GB-0006126.
XX
PA (BRBI-) BRITISH BIOTECH PHARM LTD.
XX
PI Edwards RM;
XX
DR WPI; 1997-489648/45.
DR N-PSDB; AAT85308.
XX
PT Cells for surface expression of mutants of natural G-protein
PT receptor - where the mutants can bind a ligand that the natural
PT receptor cannot, useful as biosensors
XX
PS Example 1; Pages 13-15; 35pp; English.
XX
CC The present sequence is the Saccharomyces cerevisiae alpha
CC mating factor receptor Ste2, which was used in the development of
CC a method to prepare cells presenting on their surface a mutant
CC G-protein coupled receptor (GPCR) binding a desired ligand which
CC the natural receptor cannot. The method comprises generating
CC replicable expression plasmids comprising a coding sequence for a
CC GPCR mutated by one or more amino acids in its transmembrane and/or
CC extracellular domains, and operatively linked expression control
CC sequences, transforming host cells with the plasmids and contacting
CC the cells with the ligand and isolating those which bind. GPCR are
CC natural receptor proteins, i.e. membrane proteins that allow cells
CC to detect molecules in the extracellular environment. Cells
CC presenting mutant GPCR on their surfaces are useful as biosensors,
CC with potential applications in diagnostics and environmental and
CC on-line process monitoring..
XX
SQ Sequence 481 AA;

Query Match 82.7%; Score 62; DB 18; Length 481;
Best Local Similarity 90.9%; Pred. No. 0.073;
Matches 10; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 3 SGYPYDVPDYA 13
:|||||||||||
Db 445 agypydvpdya 455

RESULT 12
AAW73372
ID AAW73372 standard; peptide; 10 AA.
XX
AC AAW73372;
XX
DT 12-FEB-1999 (first entry)
XX
DE Haemagglutinin epitope.
XX
KW TATA-box binding protein; epitope-tagged TBP; transcription complex; TAF;
KW TBP associated factor; TAF-interaction factor; gene expression regulator;
KW haemagglutinin.
XX
OS Influenza virus.
XX
PN EP881288-A1.
XX
PD 02-DEC-1998.
XX
PF 26-MAY-1998; 98EP-0109516.
XX
PR 26-MAY-1997; 97EP-0108433.
XX
PA (FARH) HOECHST AG.
XX
PI Berglund E, Kirschbaum B, Meisterernst M, Polites G;
XX
DR WPI; 1999-001394/01.
XX

PT Transgenic animal expressing epitope-tagged TATA-box binding protein
PT - for isolating higher-order transcription complexes and specific
PT factors that associate with the protein, useful as potential
PT therapeutic agents
XX
PS Claim 17; Page 18; 38pp; English.
XX
CC This sequence represents a haemagglutinin epitope. It was used in an
CC epitope-tagged TATA-box binding protein (TBP) that is expressed by the
CC transgenic non-human animals of the invention. The animals are used to
CC produce TBP. TBP is used to isolate and characterise higher-order
CC transcription complexes (from different tissue and cell types, optionally
CC at different developmental stages). It is also used to identify new
CC and/or specific TBP associated factors (TAFs, e.g. transcription factors,
CC activators or inhibitors) and TAF-interaction factors, and to raise
CC antibodies against TBP. The TAFs may be useful for regulating gene
CC expression, e.g. disease-related genes, so are potential pharmaceuticals,
CC also for identifying human analogues for use in drug screening. The
CC antibodies are used for affinity purification of TBP and its complexes.
CC TBP can isolate transcription complexes from a wide variety of different
CC tissues and cells (contrast known methods that are limited to isolation
CC from a particular cell type).
XX
SQ Sequence 10 AA;

Query Match 81.3%; Score 61; DB 20; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.0019;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 GYPYDVPDYA 13
||||||||||
Db 1 gypydvpdya 10

RESULT 13
AAW73371
ID AAW73371 standard; peptide; 11 AA.
XX
AC AAW73371;
XX
DT 12-FEB-1999 (first entry)
XX
DE Haemagglutinin epitope.
XX
KW TATA-box binding protein; epitope-tagged TBP; transcription complex; TAF;
KW TBP associated factor; TAF-interaction factor; gene expression regulator;
KW haemagglutinin.
XX
OS Influenza virus.
XX
PN EP881288-A1.
XX
PD 02-DEC-1998.
XX
PF 26-MAY-1998; 98EP-0109516.
XX
PR 26-MAY-1997; 97EP-0108433.
XX
PA (FARH) HOECHST AG.
XX
PI Berglund E, Kirschbaum B, Meisterernst M, Polites G;
XX
DR WPI; 1999-001394/01.
XX
PT Transgenic animal expressing epitope-tagged TATA-box binding protein
PT - for isolating higher-order transcription complexes and specific
PT factors that associate with the protein, useful as potential
PT therapeutic agents
XX
PS Claim 17; Page 18; 38pp; English.
XX
CC This sequence represents a haemagglutinin epitope. It was used in an
CC epitope-tagged TATA-box binding protein (TBP) that is expressed by the
CC transgenic non-human animals of the invention. The animals are used to
CC produce TBP. TBP is used to isolate and characterise higher-order
CC transcription complexes (from different tissue and cell types, optionally
CC at different developmental stages). It is also used to identify new
CC and/or specific TBP associated factors (TAFs, e.g. transcription factors,
CC activators or inhibitors) and TAF-interaction factors, and to raise
CC antibodies against TBP. The TAFs may be useful for regulating gene
CC expression, e.g. disease-related genes, so are potential pharmaceuticals,
CC also for identifying human analogues for use in drug screening. The
CC antibodies are used for affinity purification of TBP and its complexes.
CC TBP can isolate transcription complexes from a wide variety of different
CC tissues and cells (contrast known methods that are limited to isolation
CC from a particular cell type).
XX
SQ Sequence 11 AA;

Query Match 81.3%; Score 61; DB 20; Length 11;
Best Local Similarity 100.0%; Pred. No. 0.0021;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 GYPYDVPDYA 13
||||||||||
Db 2 gypydvpdya 11

RESULT 14
AAB73290
ID AAB73290 standard; Peptide; 11 AA.
XX
AC AAB73290;
XX
DT 29-MAY-2001 (first entry)
XX
DE Haemagglutinin, HA, peptide tag.
XX
KW CREB; HA peptide tag; transgenic mouse; transcription factor;
KW congestive heart failure; CHF; haemagglutinin.
XX
OS Influenza virus.
XX
PN US6194632-B1.
XX
PD 27-FEB-2001.
XX
PF 18-DEC-1998; 98US-0215098.
XX
PR 18-DEC-1997; 97US-0068011.
XX
PA (LEID/) LEIDEN J M.
XX
PI Leiden JM;
XX
DR WPI; 2001-256289/26.
XX
PT Transgenic mice expressing CREB, useful as genetic models for
PT congestive heart failure, e.g. progressive biventricular failure,
PT cardiac dilation, decreased myocardial contractility, peripheral edema
PT or intracardiac thrombi -
XX
PS Disclosure; Fig 5; 19pp; English.
XX
CC The present invention relates to transgenic mice comprising a transgene
CC under operational control of a myocyte-specific promoter. The transgene
CC encodes a protein having wild-type CREB activity or dominant negative
CC CREB activity. CREB is a 43 kD basic leucine zipper transcription factor.
CC The expression of the transgene leads to congestive heart failure (CHF).
CC The transgenic mice are useful as genetic models of dilated
CC cardiomyopathy or phenotypes associated with CHF. The transgenic mice are
CC also useful for providing important basic information concerning the role
CC of the CREB transcriptional pathway in regulating cardiac myocyte
CC function. These animals will also facilitate studies designed to identify

```
CC   new drugs that can decrease the morbidity and mortality associated with
CC   the disease. The present sequence is influenza haemagglutinin (HA)
CC   peptide tag, which was used in the present invention.
XX
SQ   Sequence   11 AA;


  Query Match             81.3%;  Score 61;  DB 22;  Length 11;
  Best Local Similarity   100.0%;  Pred. No. 0.0021;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       4 GYPYDVPDYA 13
           ||||||||||
Db       2 gypydvpdya 11


RESULT  15
AAW73370
ID   AAW73370 standard; peptide; 12 AA.
XX
AC   AAW73370;
XX
DT   12-FEB-1999  (first entry)
XX
DE   Haemagglutinin epitope.
XX
KW   TATA-box binding protein; epitope-tagged TBP; transcription complex; TAF;
KW   TBP associated factor; TAF-interaction factor; gene expression regulator;
KW   haemagglutinin.
XX
OS   Influenza virus.
XX
PN   EP881288-A1.
XX
PD   02-DEC-1998.
XX
PF   26-MAY-1998;   98EP-0109516.
XX
PR   26-MAY-1997;   97EP-0108433.
XX
PA   (FARH ) HOECHST AG.
XX
PI   Berglund E,  Kirschbaum B,  Meisterernst M,  Polites G;
XX
DR   WPI; 1999-001394/01.
XX
PT   Transgenic animal expressing epitope-tagged TATA-box binding protein
PT   - for isolating higher-order transcription complexes and specific
PT   factors that associate with the protein, useful as potential
PT   therapeutic agents
XX
PS   Claim 17; Page 16; 38pp; English.
XX
CC   This sequence represents a haemagglutinin epitope. It was used in an
CC   epitope-tagged TATA-box binding protein (TBP) that is expressed by the
CC   transgenic non-human animals of the invention. The animals are used to
CC   produce TBP. TBP is used to isolate and characterise higher-order
CC   transcription complexes (from different tissue and cell types, optionally
CC   at different developmental stages). It is also used to identify new
CC   and/or specific TBP associated factors (TAFs, e.g. transcription factors,
CC   activators or inhibitors) and TAF-interaction factors, and to raise
CC   antibodies against TBP. The TAFs may be useful for regulating gene
CC   expression, e.g. disease-related genes, so are potential pharmaceuticals,
CC   also for identifying human analogues for use in drug screening. The
CC   antibodies are used for affinity purification of TBP and its complexes.
CC   TBP can isolate transcription complexes from a wide variety of different
CC   tissues and cells (contrast known methods that are limited to isolation
CC   from a particular cell type).
XX
SQ   Sequence   12 AA;


  Query Match             81.3%;  Score 61;  DB 20;  Length 12;
  Best Local Similarity   100.0%;  Pred. No. 0.0023;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       4 GYPYDVPDYA 13
           ||||||||||
Db       2 gypydvpdya 11


Search completed: April 11, 2002, 10:15:44
Job time: 94 sec
```

GenCore version 4.5

OM protein - protein search, using sw model

Run on: April 11, 2002, 10:14:44 ; Search time 24.73 Seconds
(without alignments)
11.829 Million cell updates/sec

Title: US-09-284-787-3
Perfect score: 75
Sequence: 1 SGSGYPYDVPDYA 13

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 212252 seqs, 22503292 residues

Total number of hits satisfying chosen parameters: 212252

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	67	89.3	17	4	US-08-789-333F-51	Sequence 51, Appl
2	67	89.3	19	4	US-09-169-015-52	Sequence 52, Appl
3	65	86.7	16	4	US-09-208-966-15	Sequence 15, Appl
4	61	81.3	11	4	US-09-215-098-4	Sequence 4, Appli
5	61	81.3	14	2	US-08-632-514C-29	Sequence 29, Appl
6	61	81.3	14	3	US-09-188-177-29	Sequence 29, Appl
7	59	78.7	25	1	US-08-634-060-27	Sequence 27, Appl
8	59	78.7	25	2	US-08-700-846-10	Sequence 10, Appl
9	59	78.7	27	4	US-08-672-213-39	Sequence 39, Appl
10	59	78.7	40	4	US-08-672-213-29	Sequence 29, Appl
11	58	77.3	417	1	US-08-199-780-1	Sequence 1, Appli
12	58	77.3	417	2	US-08-316-650-1	Sequence 1, Appli
13	58	77.3	417	5	PCT-US95-02251-1	Sequence 1, Appli
14	56	74.7	32	2	US-09-066-074-7	Sequence 7, Appli
15	56	74.7	32	2	US-08-555-912A-7	Sequence 7, Appli
16	56	74.7	107	1	US-08-625-209A-26	Sequence 26, Appl
17	56	74.7	107	3	US-08-853-733B-26	Sequence 26, Appl
18	56	74.7	115	1	US-08-461-859-34	Sequence 34, Appl
19	56	74.7	219	2	US-08-902-516-2	Sequence 2, Appli
20	55	73.3	9	1	US-08-011-398B-6	Sequence 6, Appli
21	55	73.3	9	1	US-08-090-148-2	Sequence 2, Appli
22	55	73.3	9	1	US-08-370-225-6	Sequence 6, Appli
23	55	73.3	9	1	US-08-167-982-1	Sequence 1, Appli
24	55	73.3	9	1	US-08-634-060-31	Sequence 31, Appl
25	55	73.3	9	1	US-08-464-051-6	Sequence 6, Appli
26	55	73.3	9	1	US-08-461-859-6	Sequence 6, Appli
27	55	73.3	9	1	US-08-470-837-28	Sequence 28, Appl
28	55	73.3	9	2	US-08-478-386A-19	Sequence 19, Appl
29	55	73.3	9	2	US-08-292-597-19	Sequence 19, Appl
30	55	73.3	9	2	US-08-462-498-6	Sequence 6, Appli
31	55	73.3	9	2	US-08-922-267A-60	Sequence 60, Appl
32	55	73.3	9	2	US-08-388-653-19	Sequence 19, Appl
33	55	73.3	9	2	US-08-473-985-19	Sequence 19, Appl
34	55	73.3	9	2	US-08-690-011A-17	Sequence 17, Appl
35	55	73.3	9	2	US-08-690-011A-49	Sequence 49, Appl
36	55	73.3	9	2	US-08-685-625A-4	Sequence 4, Appli
37	55	73.3	9	2	US-08-687-559-14	Sequence 14, Appl
38	55	73.3	9	2	US-08-700-846-11	Sequence 11, Appl
39	55	73.3	9	2	US-08-287-537B-3	Sequence 3, Appli
40	55	73.3	9	2	US-08-483-898-19	Sequence 19, Appl
41	55	73.3	9	3	US-09-087-716-19	Sequence 19, Appl
42	55	73.3	9	3	US-08-853-733B-28	Sequence 28, Appl
43	55	73.3	9	3	US-08-554-385-6	Sequence 6, Appli
44	55	73.3	9	3	US-08-518-835-6	Sequence 6, Appli
45	55	73.3	9	3	US-09-157-753-19	Sequence 19, Appl

ALIGNMENTS

```
RESULT   1
US-08-789-333F-51
; Sequence 51, Application US/08789333F
; Patent No. 6153380
; GENERAL INFORMATION:
;  APPLICANT:  No. 6153380an, Garry P
;  APPLICANT:  Rothenberg, S. M.
;  TITLE OF INVENTION:  METHODS FOR SCREENING FOR TRANSDOMINANT INTRACELLULAR
;  TITLE OF INVENTION:  EFFECTOR PEPTIDES AND RNA MOLECULES
;  FILE REFERENCE:  A642601DJBRMSDSS
;  CURRENT APPLICATION NUMBER: US/08/789,333F
;  CURRENT FILING DATE:  1997-01-23
;  PRIOR APPLICATION NUMBER:  08/589,108
;  PRIOR FILING DATE:  1996-01-23
;  PRIOR APPLICATION NUMBER:  08/589,911
;  PRIOR FILING DATE:  1996-01-23
;  NUMBER OF SEQ ID NOS:  102
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 51
;   LENGTH: 17
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: epitope tag
;   OTHER INFORMATION: sequence.
US-08-789-333F-51


  Query Match             89.3%;  Score 67;  DB 4;  Length 17;
  Best Local Similarity   91.7%;  Pred. No. 0.00011;
  Matches   11;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       2 GSGYPYDVPDYA 13
           | ||||||||||
Db       2 GGGYPYDVPDYA 13


RESULT   2
US-09-169-015-52
; Sequence 52, Application US/09169015
; Patent No. 6180343
; GENERAL INFORMATION:
;  APPLICANT:  Anderson, David
;  APPLICANT:  Bogenberger, Jakob M.
;  TITLE OF INVENTION:  Green Fluorescent Protein Fusions with Random Peptides
;  FILE REFERENCE:  A66900/DJB/RMS/SJR
;  CURRENT APPLICATION NUMBER: US/09/169,015
;  CURRENT FILING DATE:  1998-10-08
;  NUMBER OF SEQ ID NOS:  55
;  SOFTWARE:  PatentIn Ver. 2.0
```

```
; SEQ ID NO 52
;   LENGTH: 19
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:  synthetic
US-09-169-015-52


  Query Match             89.3%;  Score 67;  DB 4;  Length 19;
  Best Local Similarity   91.7%;  Pred. No. 0.00013;
  Matches   11;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      2 GSGYPYDVPDYA 13
          | ||||||||||
Db      2 GGGYPYDVPDYA 13


RESULT   3
US-09-208-966-15
; Sequence 15, Application US/09208966
; Patent No. 6221355
; GENERAL INFORMATION:
;  APPLICANT: Dowdy, Steven F.
;  TITLE OF INVENTION: ANTI-PATHOGEN SYSTEM AND METHODS OF USE THEREOF
;  FILE REFERENCE: 48881/1742
;  CURRENT APPLICATION NUMBER: US/09/208,966
;  CURRENT FILING DATE: 1998-12-10
;  EARLIER APPLICATION NUMBER: 60/082,402
;  EARLIER FILING DATE: 1998-04-20
;  EARLIER APPLICATION NUMBER: 60/069,012
;  EARLIER FILING DATE: 1997-12-10
;  NUMBER OF SEQ ID NOS: 57
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 15
;   LENGTH: 16
;   TYPE: PRT
;   ORGANISM: human
US-09-208-966-15


  Query Match             86.7%;  Score 65;  DB 4;  Length 16;
  Best Local Similarity  100.0%;  Pred. No. 0.00022;
  Matches   11;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      3 SGYPYDVPDYA 13
          ||||||||||||
Db      2 SGYPYDVPDYA 12


RESULT   4
US-09-215-098-4
; Sequence 4, Application US/09215098
; Patent No. 6194632
; GENERAL INFORMATION:
;  APPLICANT:  Leiden, Jeffery M
;  TITLE OF INVENTION:  DILATED CARDIOMYOPATHY IN TRANSGENIC MICE EXPRESSING A
;  TITLE OF INVENTION:  DOMINANT-NEGATIVE CREB TRANSCRIPTION FACTOR IN THE
;  TITLE OF INVENTION:  HEART
;  FILE REFERENCE:  9189-4
;  CURRENT APPLICATION NUMBER: US/09/215,098
;  CURRENT FILING DATE:  1998-12-18
;  PRIOR APPLICATION NUMBER:  60/068,011
;  PRIOR FILING DATE:  1997-12-18
;  NUMBER OF SEQ ID NOS:  5
;  SOFTWARE:  PatentIn Ver. 2.1
; SEQ ID NO 4
;   LENGTH: 11
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:HA peptide tag
US-09-215-098-4


  Query Match             81.3%;  Score 61;  DB 4;  Length 11;
  Best Local Similarity  100.0%;  Pred. No. 0.00062;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      4 GYPYDVPDYA 13
          ||||||||||
Db      2 GYPYDVPDYA 11


RESULT   5
US-08-632-514C-29
; Sequence 29, Application US/08632514C
; Patent No. 5834234
;  GENERAL INFORMATION:
;    APPLICANT:  GALLO, Gregory J.
;    TITLE OF INVENTION:  APOPTOSIS ASSOCIATED PROTEIN BBK
;    NUMBER OF SEQUENCES:  29
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Hale and Dorr
;      STREET:  1455 Pennsylvania Avenue, N.W.
;      CITY:  Washington, D.C.
;      COUNTRY:  U.S.A.
;      ZIP:  20004
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/632,514C
;      FILING DATE:  29-MAY-1996
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  WIXON, Henry N.
;      REGISTRATION NUMBER:  32,073
;      REFERENCE/DOCKET NUMBER:  104322.188
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202)942-8459
;      TELEFAX:  (202)942-8484
;  INFORMATION FOR SEQ ID NO:  29:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  14 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-632-514C-29


  Query Match             81.3%;  Score 61;  DB 2;  Length 14;
  Best Local Similarity  100.0%;  Pred. No. 0.0008;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      4 GYPYDVPDYA 13
          ||||||||||
Db      2 GYPYDVPDYA 11


RESULT   6
US-09-188-177-29
; Sequence 29, Application US/09188177
; Patent No. 6057132
;  GENERAL INFORMATION:
;    APPLICANT:  GALLO, Gregory J.
;    TITLE OF INVENTION:  APOPTOSIS ASSOCIATED PROTEIN BBK
;    NUMBER OF SEQUENCES:  29
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Hale and Dorr
;      STREET:  1455 Pennsylvania Avenue, N.W.
;      CITY:  Washington, D.C.
;      COUNTRY:  U.S.A.
```

```
;      ZIP:  20004
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/188,177
;      FILING DATE:  2
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  WIXON, Henry N.
;      REGISTRATION NUMBER:  32,073
;      REFERENCE/DOCKET NUMBER:  104322.188
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202)942-8459
;      TELEFAX:  (202)942-8484
;  INFORMATION FOR SEQ ID NO:  29:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  14 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-09-188-177-29


  Query Match             81.3%;  Score 61;  DB 3;  Length 14;
  Best Local Similarity   100.0%;  Pred. No. 0.0008;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      4 GYPYDVPDYA 13
          ||||||||||
Db      2 GYPYDVPDYA 11


RESULT   7
US-08-634-060-27
; Sequence 27, Application US/08634060
; Patent No. 5712136
;  GENERAL INFORMATION:
;    APPLICANT:  Wickham, Thomas J.
;    APPLICANT:  Kovesdi, Imre
;    APPLICANT:  Roelvink, Petrus W.
;    TITLE OF INVENTION:  ADENOVIRAL-MEDIATED CELL TARGETING COMMANDED BY
;    TITLE OF INVENTION:  THE ADENOVIRUS PENTON BASE PROTEIN
;    NUMBER OF SEQUENCES:  60
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Leydig, Voit & Mayer, Ltd.
;      STREET:  Two Prudential Plaza, Suite 4900
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  USA
;      ZIP:  60601
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/634,060
;      FILING DATE:
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/303,162
;      FILING DATE:  08-SEP-1994
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Kilyk, John Jr.
;      REGISTRATION NUMBER:  30763
;      REFERENCE/DOCKET NUMBER:  71602
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (312) 616-5600
;      TELEFAX:  (312) 616-5700
;  INFORMATION FOR SEQ ID NO:  27:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  25 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-634-060-27


  Query Match             78.7%;  Score 59;  DB 1;  Length 25;
  Best Local Similarity   83.3%;  Pred. No. 0.0031;
  Matches   10;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      2 GSGYPYDVPDYA 13
            |  |||||||||
Db     11 GDTYPYDVPDYA 22


RESULT   8
US-08-700-846-10
; Sequence 10, Application US/08700846
; Patent No. 5962311
;  GENERAL INFORMATION:
;    APPLICANT:  WICKHAM, THOMAS J.
;    APPLICANT:  ROELVINK, PETRUS W.
;    APPLICANT:  KOVESDI, IMRE
;    TITLE OF INVENTION:  A SHORT-SHAFTED ADENOVIRAL FIBER AND ITS
;    TITLE OF INVENTION:  USE
;    NUMBER OF SEQUENCES:  17
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  LEYDIG, VOIT & MAYER, LTD.
;      STREET:  TWO PRUDENTIAL PLAZA, SUITE 4900
;      CITY:  CHICAGO
;      STATE:  IL
;      COUNTRY:  USA
;      ZIP:  60601-6780
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/700,846
;      FILING DATE:  21-AUG-1996
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  LARCHER, CAROL
;      REGISTRATION NUMBER:  35243
;      REFERENCE/DOCKET NUMBER:  74294
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (312) 616-5600
;      TELEFAX:  (312) 616-5700
;  INFORMATION FOR SEQ ID NO:  10:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  25 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-700-846-10


  Query Match             78.7%;  Score 59;  DB 2;  Length 25;
  Best Local Similarity   83.3%;  Pred. No. 0.0031;
  Matches   10;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      2 GSGYPYDVPDYA 13
            |  |||||||||
Db     11 GDTYPYDVPDYA 22
```

```
RESULT   9
US-08-672-213-39
; Sequence 39, Application US/08672213
; Patent No. 6306649
;  GENERAL INFORMATION:
;    APPLICANT:  GILMAN, Michael Z.
;    APPLICANT:  NATESAN, Sridaran
;    TITLE OF INVENTION:  USE OF HETEROLOGOUS TRANSCRIPTION
;    TITLE OF INVENTION:  FACTORS IN GENE THERAPY
;    NUMBER OF SEQUENCES:  72
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  ARIAD Gene Therapeutics, Inc.
;      STREET:  26 Landsdowne Street
;      CITY:  Cambridge
;      STATE:  Massachusetts
;      COUNTRY:  USA
;      ZIP:  02139-4234
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/672,213
;      FILING DATE:  27-JUN-1996
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/000,553
;      FILING DATE:  27-JUN-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/019,614
;      FILING DATE:  29-DEC-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  BERSTEIN, David L.
;      REGISTRATION NUMBER:  31,235
;      REFERENCE/DOCKET NUMBER:  ARIAD 346B
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617-494-0400
;      TELEFAX:  617-494-0208
;  INFORMATION FOR SEQ ID NO:  39:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  27 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-672-213-39


  Query Match             78.7%;  Score 59;  DB 4;  Length 27;
  Best Local Similarity   90.9%;  Pred. No. 0.0034;
  Matches   10;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       3 SGYPYDVPDYA 13
           | |||||||||
Db       3 SSYPYDVPDYA 13


RESULT  10
US-08-672-213-29
; Sequence 29, Application US/08672213
; Patent No. 6306649
;  GENERAL INFORMATION:
;    APPLICANT:  GILMAN, Michael Z.
;    APPLICANT:  NATESAN, Sridaran
;    TITLE OF INVENTION:  USE OF HETEROLOGOUS TRANSCRIPTION
;    TITLE OF INVENTION:  FACTORS IN GENE THERAPY
;    NUMBER OF SEQUENCES:  72
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  ARIAD Gene Therapeutics, Inc.
;      STREET:  26 Landsdowne Street
;      CITY:  Cambridge
;      STATE:  Massachusetts
;      COUNTRY:  USA
;      ZIP:  02139-4234
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/672,213
;      FILING DATE:  27-JUN-1996
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/000,553
;      FILING DATE:  27-JUN-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/019,614
;      FILING DATE:  29-DEC-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  BERSTEIN, David L.
;      REGISTRATION NUMBER:  31,235
;      REFERENCE/DOCKET NUMBER:  ARIAD 346B
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617-494-0400
;      TELEFAX:  617-494-0208
;  INFORMATION FOR SEQ ID NO:  29:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  40 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-672-213-29


  Query Match             78.7%;  Score 59;  DB 4;  Length 40;
  Best Local Similarity   90.9%;  Pred. No. 0.0052;
  Matches   10;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       3 SGYPYDVPDYA 13
           | |||||||||
Db       3 SSYPYDVPDYA 13


RESULT  11
US-08-199-780-1
; Sequence 1, Application US/08199780
; Patent No. 5763416
;  GENERAL INFORMATION:
;    APPLICANT:  Bonadio, Jeffrey
;    APPLICANT:  Goldstein, Steven A.
;    TITLE OF INVENTION:  Gene Transfer Into Bone Cells
;    TITLE OF INVENTION:  And Tissues
;    NUMBER OF SEQUENCES:  3
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Arnold, White & Durkee
;      STREET:  P.O. Box 4433
;      CITY:  Houston
;      STATE:  TX
;      COUNTRY:  USA
;      ZIP:  77210
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30B
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/199,780
;      FILING DATE:  18-FEB-1994
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Parker, David
;      REGISTRATION NUMBER:  32,165
;      REFERENCE/DOCKET NUMBER:  UMIC:002
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (512) 320-7200
```

```
;      TELEFAX:  (512) 474-7577
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  417 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-199-780-1


  Query Match           77.3%;  Score 58;  DB 1;  Length 417;
  Best Local Similarity  83.3%;  Pred. No. 0.095;
  Matches   10;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       2 GSGYPYDVPDYA 13
           |  |||||||||
Db     406 GCRYPYDVPDYA 417


RESULT  12
US-08-316-650-1
; Sequence 1, Application US/08316650
; Patent No. 5942496
;  GENERAL INFORMATION:
;    APPLICANT:  Bonadio, Jeffrey
;    APPLICANT:  Roessler, Blake J.
;    APPLICANT:  Goldstein, Steven A.
;    APPLICANT:  Lin, Wushan
;    TITLE OF INVENTION:  METHODS AND COMPOSITIONS
;    TITLE OF INVENTION:  FOR STIMULATING BONE CELLS
;    NUMBER OF SEQUENCES:  15
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Arnold, White & Durkee
;      STREET:  P.O. Box 4433
;      CITY:  Houston
;      STATE:  Texas
;      COUNTRY:  USA
;      ZIP:  77210
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/316,650
;      FILING DATE:  30-SEP-1994
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/199,780
;      FILING DATE:  30-SEP-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Parker, David L.
;      REGISTRATION NUMBER:  32,165
;      REFERENCE/DOCKET NUMBER:  UMIC:008
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (512) 418-3000
;      TELEFAX:  (713) 789-2679
;      TELEX:  79-0924
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  417 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-316-650-1


  Query Match           77.3%;  Score 58;  DB 2;  Length 417;
  Best Local Similarity  83.3%;  Pred. No. 0.095;
  Matches   10;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps   .0;

Qy       2 GSGYPYDVPDYA 13
           |  |||||||||
Db     406 GCRYPYDVPDYA 417


RESULT  13
PCT-US95-02251-1
; Sequence 1, Application PC/TUS9502251
;  GENERAL INFORMATION:
;    APPLICANT:
;    TITLE OF INVENTION:  METHODS AND COMPOSITIONS FOR STIMULATING BONE
;    TITLE OF INVENTION:  CELLS
;    NUMBER OF SEQUENCES:  18
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Arnold, White & Durkee
;      STREET:  P.O. Box 4433
;      CITY:  Houston
;      STATE:  Texas
;      COUNTRY:  United States of America
;      ZIP:  77210
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS/ASCII
;      SOFTWARE:  PatentIn Release #1.0, Version
;      SOFTWARE:  #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US95/02251
;      FILING DATE:  CONCURRENTLY HEREWITH
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/316,650
;      FILING DATE:  30-SEP-1994
;      CLASSIFICATION:
;      APPLICATION NUMBER:  US 08/199,780
;      FILING DATE:  18-FEB-1994
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Parker, David L.
;      REGISTRATION NUMBER:  32,165
;      REFERENCE/DOCKET NUMBER:  UMIC009P--
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (512) 418-3000
;      TELEFAX:  (713) 789-2679
;      TELEX:  79-0924
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  417 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
PCT-US95-02251-1


  Query Match           77.3%;  Score 58;  DB 5;  Length 417;
  Best Local Similarity  83.3%;  Pred. No. 0.095;
  Matches   10;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       2 GSGYPYDVPDYA 13
           |  |||||||||
Db     406 GCRYPYDVPDYA 417


RESULT  14
US-09-066-074-7
; Sequence 7, Application US/09066074
; Patent No. 5952467
;  GENERAL INFORMATION:
;    APPLICANT:  Hunter et al., Tony
;    TITLE OF INVENTION:  NIMA INTERACTING PROTEINS
;    NUMBER OF SEQUENCES:  14
;    CORRESPONDENCE ADDRESS:
```

```
;      ADDRESSEE:  Fish & Richardson P.C.
;      STREET:  4225 Executive Square, Suite 1400
;      CITY:  La Jolla
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  92037
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/066,074
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/555,912
;      FILING DATE:  13-NOV-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Haile, Lisa A.
;      REGISTRATION NUMBER:  38,347
;      REFERENCE/DOCKET NUMBER:  07251/011001
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  619/678-5070
;      TELEFAX:  619/678-5099
;  INFORMATION FOR SEQ ID NO:  7:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  32 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  not relevant
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    IMMEDIATE SOURCE:
;      CLONE:  PGAL-PIN1
US-09-066-074-7


  Query Match             74.7%;  Score 56;  DB 2;  Length 32;
  Best Local Similarity   81.8%;  Pred. No. 0.012;
  Matches    9;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       3 SGYPYDVPDYA 13
           : |||||||||
Db       2 ASYPYDVPDYA 12


RESULT  15
US-08-555-912A-7
; Sequence 7, Application US/08555912A
; Patent No. 5972697
;  GENERAL INFORMATION:
;    APPLICANT:  Hunter et al., Tony
;    TITLE OF INVENTION:  NIMA INTERACTING PROTEINS
;    NUMBER OF SEQUENCES:  14
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Fish & Richardson P.C.
;      STREET:  4225 Executive Square, Suite 1400
;      CITY:  La Jolla
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  92037
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/555,912A
;      FILING DATE:  13-NOV-1995
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Haile, Lisa A.
;      REGISTRATION NUMBER:  38,347
;      REFERENCE/DOCKET NUMBER:  07251/011001
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  619/678-5070
;      TELEFAX:  619/678-5099
;  INFORMATION FOR SEQ ID NO:  7:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  32 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  not relevant
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    IMMEDIATE SOURCE:
;      CLONE:  PGAL-PIN1
US-08-555-912A-7


  Query Match             74.7%;  Score 56;  DB 2;  Length 32;
  Best Local Similarity   81.8%;  Pred. No. 0.012;
  Matches    9;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       3 SGYPYDVPDYA 13
           : |||||||||
Db       2 ASYPYDVPDYA 12


Search completed: April 11, 2002, 10:14:44
Job time: 34 sec
```

GenCore version 4.5

OM protein - protein search, using sw model

Run on: April 11, 2002, 10:16:16 ; Search time 27.06 Seconds
(without alignments)
36.595 Million cell updates/sec

Title: US-09-284-787-3
Perfect score: 75
Sequence: 1 SGSGYPYDVPDYA 13

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_68:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	56	74.7	347	2	S52188	hemagglutinin - in
2	56	74.7	347	2	S52191	hemagglutinin - in
3	56	74.7	550	1	HMIVS2	hemagglutinin prec
4	56	74.7	550	2	JQ1153	hemagglutinin prec
5	56	74.7	565	1	HMIVET	hemagglutinin prec
6	55	73.3	330	2	JQ2374	hemagglutinin - in
7	55	73.3	330	2	JQ2375	hemagglutinin - in
8	55	73.3	331	2	JQ2377	hemagglutinin - in
9	55	73.3	331	2	JQ2378	hemagglutinin - in
10	55	73.3	347	2	S52173	hemagglutinin - in
11	55	73.3	347	2	S52174	hemagglutinin - in
12	55	73.3	347	2	S52175	hemagglutinin - in
13	55	73.3	347	2	S52176	hemagglutinin - in
14	55	73.3	347	2	S52178	hemagglutinin - in
15	55	73.3	347	2	S52179	hemagglutinin - in
16	55	73.3	347	2	S52180	hemagglutinin - in
17	55	73.3	347	2	S52182	hemagglutinin - in
18	55	73.3	347	2	S52183	hemagglutinin - in
19	55	73.3	347	2	S52185	hemagglutinin - in
20	55	73.3	347	2	S52186	hemagglutinin - in
21	55	73.3	347	2	S52187	hemagglutinin - in
22	55	73.3	347	2	S52189	hemagglutinin - in
23	55	73.3	347	2	S52195	hemagglutinin - in
24	55	73.3	347	2	S52197	hemagglutinin - in
25	55	73.3	347	2	S52199	hemagglutinin - in
26	55	73.3	362	2	S38637	hemagglutinin - in
27	55	73.3	550	1	HMIVS3	hemagglutinin prec
28	55	73.3	550	1	HMIV77	hemagglutinin prec
29	55	73.3	550	1	HMIV80	hemagglutinin prec
30	55	73.3	550	1	HMIV33	hemagglutinin prec
31	55	73.3	550	1	HMIV89	hemagglutinin prec
32	55	73.3	550	1	HMIV98	hemagglutinin prec
33	55	73.3	550	1	HMIV15	hemagglutinin prec
34	55	73.3	550	1	HMIV86	hemagglutinin prec
35	55	73.3	550	2	JQ1154	hemagglutinin prec
36	55	73.3	550	2	JQ1155	hemagglutinin prec
37	55	73.3	550	2	JQ1156	hemagglutinin prec
38	55	73.3	565	1	HMIVE2	hemagglutinin prec
39	55	73.3	565	1	HMIVE4	hemagglutinin prec
40	55	73.3	566	1	HMIVH	hemagglutinin prec
41	55	73.3	566	1	HMIVHA	hemagglutinin prec
42	55	73.3	566	1	HMIVHM	hemagglutinin prec
43	55	73.3	566	1	HMIV6	hemagglutinin prec
44	55	73.3	567	1	HMIVV	hemagglutinin prec
45	54	72.0	565	1	HMIVE1	hemagglutinin prec

ALIGNMENTS

RESULT 1
S52188
hemagglutinin - influenza A virus (isolate A/Guangdong/25/93) (fragment)
C;Species: influenza A virus
A;Variety: isolate A/Guangdong/25/93
C;Date: 07-May-1995 #sequence_revision 23-Feb-1996 #text_change 23-Mar-2001
C;Accession: S52188
R;Ellis, J.S.
submitted to the EMBL Data Library, October 1994
A;Description: Antigenic and genetic variation in the haemagglutinin of recently circ
A;Reference number: S52173
A;Accession: S52188
A;Molecule type: mRNA
A;Residues: 1-347 <ELL>
A;Cross-references: EMBL:Z46406; NID:g609049; PIDN:CAA86541.1; PID:g609050
A;Experimental source: isolate A/Guangdong/25/93
C;Superfamily: influenza virus hemagglutinin

Query Match 74.7%; Score 56; DB 2; Length 347;
Best Local Similarity 90.9%; Pred. No. 0.17;
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy       3 SGYPYDVPDYA 13
           | |||||||||
Db      96 SCYPYDVPDYA 106
```

RESULT 2
S52191
hemagglutinin - influenza A virus (isolate A/Hong Kong/2/94) (fragment)
C;Species: influenza A virus
A;Variety: isolate A/Hong Kong/2/94
C;Date: 07-May-1995 #sequence_revision 23-Feb-1996 #text_change 23-Mar-2001
C;Accession: S52191
R;Ellis, J.S.
submitted to the EMBL Data Library, October 1994
A;Description: Antigenic and genetic variation in the haemagglutinin of recently circ
A;Reference number: S52173
A;Accession: S52191
A;Molecule type: mRNA
A;Residues: 1-347 <ELL>
A;Cross-references: EMBL:Z46408; NID:g609055; PIDN:CAA86543.1; PID:g609056
A;Experimental source: isolate A/Hong Kong/2/94
C;Superfamily: influenza virus hemagglutinin

Query Match 74.7%; Score 56; DB 2; Length 347;
Best Local Similarity 90.9%; Pred. No. 0.17;
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy       3 SGYPYDVPDYA 13
```

```
          | |||||||||
Db      96 SCYPYDVPDYA 106


RESULT   3
HMIVS2
hemagglutinin precursor - influenza A virus (strain A/swine/126/82) (fragment)
C;Species: influenza A virus
C;Date: 31-Mar-1989 #sequence_revision 31-Mar-1989 #text_change 18-Sep-1998
C;Accession: A29971
R;Kida, H.; Shortridge, K.F.; Webster, R.G.
Virology 162, 160-166, 1988
A;Title: Origin of the hemagglutinin gene of H3N2 influenza viruses from pigs in China.
A;Reference number: A94370; MUID:88101364
A;Accession: A29971
A;Molecule type: genomic RNA
A;Residues: 1-550 <KID>
A;Cross-references: GB:M19056; NID:g324208
A;Note: the sequence in GenBank entry FLAHAPA, release 106, (PID:g324209) differs from t
C;Genetics:
A;Map position: segment 4
C;Superfamily: influenza virus hemagglutinin
C;Keywords: glycoprotein; hemagglutinin; homotrimer; lipoprotein; thiolester bond
F;1-328/Product: hemagglutinin HA1 #status predicted <HA1>
F;330-550/Product: hemagglutinin HA2 #status predicted <HA2>
F;520-536/Domain: transmembrane #status predicted <TM1>
F;8,22,38,165,285,483/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;14-466,52-277,64-76,139-473,281-305/Disulfide bonds: #status predicted
F;539,546,549/Binding site: palmitate (Cys) (covalent) #status predicted


  Query Match             74.7%;  Score 56;  DB 1;  Length 550;
  Best Local Similarity   90.9%;  Pred. No. 0.27;
  Matches   10;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       3 SGYPYDVPDYA 13
           | |||||||||
Db      96 SCYPYDVPDYA 106


RESULT   4
JQ1153
hemagglutinin precursor - influenza A virus (strain A/duck/Hong Kong/7/75) (fragment)
N;Contains: hemagglutinin HA1; hemagglutinin HA2
C;Species: influenza A virus
C;Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 20-Jun-2000
C;Accession: JQ1153
R;Yasuda, J.; Shortridge, K.F.; Shimizu, Y.; Kida, H.
J. Gen. Virol. 72, 2007-2010, 1991
A;Title: Molecular evidence for a role of domestic ducks in the introduction of avian H3
A;Reference number: JQ1153; MUID:91341491
A;Accession: JQ1153
A;Molecule type: genomic RNA
A;Residues: 1-550 <YAS>
A;Cross-references: GB:D00929; NID:g221279; PIDN:BAA00769.1; PID:g221280
A;Note: the authors translated the codon GGG for residue 218 as Glu
A;Note: residues 528-532 are not shown in this publication
C;Superfamily: influenza virus hemagglutinin
C;Keywords: glycoprotein; homotrimer
F;1-328/Product: hemagglutinin HA1 #status predicted <HA1>
F;330-545/Product: hemagglutinin HA2 #status predicted <HA2>
F;8,22,38,165,285,483/Binding site: carbohydrate (Asn) (covalent) #status predicted


  Query Match             74.7%;  Score 56;  DB 2;  Length 550;
  Best Local Similarity   90.9%;  Pred. No. 0.27;
  Matches   10;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       3 SGYPYDVPDYA 13
           | |||||||||
Db      96 SCYPYDVPDYA 106


RESULT   5
HMIVET
hemagglutinin precursor - influenza A virus (strain A/equine/Kentucky/2/86[H3N8])
N;Contains: hemagglutinin HA1; hemagglutinin HA2
C;Species: influenza A virus
C;Date: 30-Sep-1990 #sequence_revision 30-Sep-1990 #text_change 16-Jul-1999
C;Accession: A34065
R;Kawaoka, Y.; Bean, W.J.; Webster, R.G.
Virology 169, 283-292, 1989
A;Title: Evolution of the hemagglutinin of equine H3 influenza viruses.
A;Reference number: A34064; MUID:89204899
A;Accession: A34065
A;Molecule type: genomic RNA
A;Residues: 1-565 <KAW>
A;Cross-references: GB:M24727; GB:J04336; NID:g324000; PIDN:AAA43102.1; PID:g324001
C;Genetics:
A;Map position: segment 4
C;Superfamily: influenza virus hemagglutinin
C;Keywords: glycoprotein; hemagglutinin; homotrimer; lipoprotein; thiolester bond
F;1-16/Domain: signal sequence #status predicted <SIG>
F;17-344/Product: hemagglutinin HA1 #status predicted <HA1>
F;345-565/Product: hemagglutinin HA2 #status predicted <HA2>
F;535-551/Domain: transmembrane #status predicted <TM1>
F;18,23,37,53,68,78,180,300,498/Binding site: carbohydrate (Asn) (covalent) #status p
F;29-481,67-292,79-91,154-488,296-320/Disulfide bonds: #status predicted
F;554,561,564/Binding site: palmitate (Cys) (covalent) #status predicted


  Query Match             74.7%;  Score 56;  DB 1;  Length 565;
  Best Local Similarity   69.2%;  Pred. No. 0.28;
  Matches    9;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 SGSGYPYDVPDYA 13
           | : ||||:||||
Db     109 SSNCYPYDIPDYA 121


RESULT   6
JQ2374
hemagglutinin - influenza A virus (strain SN1289)
C;Species: influenza A virus
C;Date: 14-Jul-1994 #sequence_revision 14-Jul-1994 #text_change 07-May-1999
C;Accession: JQ2374
R;Rocha, E.P.; Xu, X.; Hall, H.E.; Allen, J.R.; Regnery, H.L.; Cox, N.J.
J. Gen. Virol. 74, 2513-2518, 1993
A;Title: Comparison of 10 influenza A (H1N1 and H3N2) haemagglutinin sequences obtain
A;Reference number: JQ2369; MUID:94065682
A;Accession: JQ2374
A;Molecule type: mRNA
A;Residues: 1-330 <ROC>
A;Cross-references: GB:L20115
A;Experimental source: subtype H3N2
C;Superfamily: influenza virus hemagglutinin


  Query Match             73.3%;  Score 55;  DB 2;  Length 330;
  Best Local Similarity   100.0%;  Pred. No. 0.23;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       5 YPYDVPDYA 13
           |||||||||
Db      98 YPYDVPDYA 106


RESULT   7
JQ2375
hemagglutinin - influenza A virus (strain SN1389 and IN0190)
C;Species: influenza A virus
C;Date: 14-Jul-1994 #sequence_revision 14-Jul-1994 #text_change 07-May-1999
C;Accession: JQ2375; JQ2376
R;Rocha, E.P.; Xu, X.; Hall, H.E.; Allen, J.R.; Regnery, H.L.; Cox, N.J.
J. Gen. Virol. 74, 2513-2518, 1993
A;Title: Comparison of 10 influenza A (H1N1 and H3N2) haemagglutinin sequences obtain
```

```
A;Reference number: JQ2369; MUID:94065682
A;Accession: JQ2375
A;Molecule type: mRNA
A;Residues: 1-330 <ROC>
A;Cross-references: GB:L20119
A;Experimental source: subtype H3N2
C;Superfamily: influenza virus hemagglutinin


  Query Match             73.3%;  Score 55;  DB 2;  Length 330;
  Best Local Similarity   100.0%;  Pred. No. 0.23;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      5 YPYDVPDYA 13
          |||||||||
Db     98 YPYDVPDYA 106


RESULT   8
JQ2377
hemagglutinin - influenza A virus (strain SD0191)
C;Species: influenza A virus
C;Date: 14-Jul-1994 #sequence_revision 14-Jul-1994 #text_change 07-May-1999
C;Accession: JQ2377
R;Rocha, E.P.; Xu, X.; Hall, H.E.; Allen, J.R.; Regnery, H.L.; Cox, N.J.
J. Gen. Virol. 74, 2513-2518, 1993
A;Title: Comparison of 10 influenza A (H1N1 and H3N2) haemagglutinin sequences obtained
A;Reference number: JQ2369; MUID:94065682
A;Accession: JQ2377
A;Molecule type: mRNA
A;Residues: 1-331 <ROC>
A;Cross-references: GB:L20105
A;Experimental source: subtype H3N2
C;Superfamily: influenza virus hemagglutinin


  Query Match             73.3%;  Score 55;  DB 2;  Length 331;
  Best Local Similarity   100.0%;  Pred. No. 0.23;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      5 YPYDVPDYA 13
          |||||||||
Db     98 YPYDVPDYA 106


RESULT   9
JQ2378
hemagglutinin - influenza A virus (strain HI0191)
C;Species: influenza A virus
C;Date: 14-Jul-1994 #sequence_revision 14-Jul-1994 #text_change 07-May-1999
C;Accession: JQ2378
R;Rocha, E.P.; Xu, X.; Hall, H.E.; Allen, J.R.; Regnery, H.L.; Cox, N.J.
J. Gen. Virol. 74, 2513-2518, 1993
A;Title: Comparison of 10 influenza A (H1N1 and H3N2) haemagglutinin sequences obtained
A;Reference number: JQ2369; MUID:94065682
A;Accession: JQ2378
A;Molecule type: mRNA
A;Residues: 1-331 <ROC>
A;Cross-references: GB:L20102
A;Experimental source: subtype H3N2
C;Superfamily: influenza virus hemagglutinin


  Query Match             73.3%;  Score 55;  DB 2;  Length 331;
  Best Local Similarity   100.0%;  Pred. No. 0.23;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      5 YPYDVPDYA 13
          |||||||||
Db     98 YPYDVPDYA 106


RESULT  10
S52173
hemagglutinin - influenza A virus (fragment)
C;Species: influenza A virus
C;Date: 07-May-1995 #sequence_revision 23-Feb-1996 #text_change 23-Mar-2001
C;Accession: S52173; S52190; S52192; S52177; S52181; S52194; S52196; S52184; S52193
R;Ellis, J.S.
submitted to the EMBL Data Library, October 1994
A;Description: Antigenic and genetic variation in the haemagglutinin of recently circ
A;Reference number: S52173
A;Accession: S52173
A;Molecule type: mRNA
A;Residues: 1-347 <ELL>
A;Cross-references: EMBL:Z46392; NID:g609020; PIDN:CAA86527.1; PID:g609021
A;Experimental source: isolate A/Beijing/32/92
A;Accession: S52190
A;Molecule type: mRNA
A;Residues: 1-347 <EL2>
A;Cross-references: EMBL:Z46410; NID:g609053; PIDN:CAA86545.1; PID:g609054
A;Experimental source: isolate A/Hong Kong/23/92
A;Accession: S52192
A;Molecule type: mRNA
A;Residues: 1-347 <EL3>
A;Cross-references: EMBL:Z46409; NID:g609057; PIDN:CAA86544.1; PID:g609058
A;Experimental source: isolate A/Hong Kong/34/90
A;Accession: S52177
A;Molecule type: mRNA
A;Residues: 1-134,'K',136-156,'L',158-173,'F',175-188,'S',190-218,'F',220-225,'L',227
A;Cross-references: EMBL:Z46395; NID:g609027; PIDN:CAA86530.1; PID:g1228087
A;Experimental source: isolate A/England/269/93
A;Accession: S52181
A;Molecule type: mRNA
A;Residues: 1-134,'K',136-156,'L',158-173,'F',175-188,'S',190-218,'F',220-225,'L',227
A;Cross-references: EMBL:Z46399; NID:g609035; PIDN:CAA86534.1; PID:g609036
A;Experimental source: isolate A/England/328/93
A;Accession: S52194
A;Molecule type: mRNA
A;Residues: 1-134,'K',136-156,'L',158-173,'F',175-188,'S',190-218,'F',220-225,'L',227
A;Cross-references: EMBL:Z46414; NID:g609061; PIDN:CAA86549.1; PID:g609062
A;Experimental source: isolate A/Scotland/160/93
A;Accession: S52196
A;Molecule type: mRNA
A;Residues: 1-134,'K',136-156,'L',158-173,'F',175-188,'S',190-218,'F',220-225,'L',227
A;Cross-references: EMBL:Z46415; NID:g609065; PIDN:CAA86550.1; PID:g609066
A;Experimental source: isolate A/Scotland/174/93
A;Accession: S52184
A;Molecule type: mRNA
A;Residues: 1-74,'N',76-120,'T',122-156,'L',158-173,'F',175-188,'S',190-200,'K',202-2
A;Cross-references: EMBL:Z46402; NID:g609041; PIDN:CAA86537.1; PID:g609042
A;Experimental source: isolate A/England/471/93
A;Accession: S52193
A;Molecule type: mRNA
A;Residues: 1-74,'N',76-120,'T',122-156,'L',158-173,'F',175-188,'S',190-200,'K',202-2
A;Cross-references: EMBL:Z46413; NID:g609059; PIDN:CAA86548.1; PID:g609060
A;Experimental source: isolate A/Scotland/142/93
C;Superfamily: influenza virus hemagglutinin


  Query Match             73.3%;  Score 55;  DB 2;  Length 347;
  Best Local Similarity   100.0%;  Pred. No. 0.24;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      5 YPYDVPDYA 13
          |||||||||
Db     98 YPYDVPDYA 106


RESULT  11
S52174
hemagglutinin - influenza A virus (isolate A/Beijing/353/89) (fragment)
C;Species: influenza A virus
A;Variety: isolate A/Beijing/353/89
C;Date: 07-May-1995 #sequence_revision 23-Feb-1996 #text_change 20-Sep-1999
C;Accession: S52174
```

```
R;Ellis, J.S.
submitted to the EMBL Data Library, October 1994
A;Description: Antigenic and genetic variation in the haemagglutinin of recently circula
A;Reference number: S52173
A;Accession: S52174
A;Molecule type: mRNA
A;Residues: 1-347 <ELL>
A;Cross-references: EMBL:Z46391; NID:g609022; PIDN:CAA86526.1; PID:g609023
A;Experimental source: isolate A/Beijing/353/89
C;Superfamily: influenza virus hemagglutinin


  Query Match             73.3%;  Score 55;  DB 2;  Length 347;
  Best Local Similarity   100.0%;  Pred. No. 0.24;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       5 YPYDVPDYA 13
           |||||||||
Db      98 YPYDVPDYA 106


RESULT  12
S52175
hemagglutinin - influenza A virus (isolate A/England/1/93) (fragment)
C;Species: influenza A virus
A;Variety: isolate A/England/1/93
C;Date: 07-May-1995 #sequence_revision 23-Feb-1996 #text_change 23-Mar-2001
C;Accession: S52175
R;Ellis, J.S.
submitted to the EMBL Data Library, October 1994
A;Description: Antigenic and genetic variation in the haemagglutinin of recently circula
A;Reference number: S52173
A;Accession: S52175
A;Molecule type: mRNA
A;Residues: 1-347 <ELL>
A;Cross-references: EMBL:Z46393; NID:g609024; PIDN:CAA86528.1; PID:g940547
A;Experimental source: isolate A/England/1/93
C;Superfamily: influenza virus hemagglutinin


  Query Match             73.3%;  Score 55;  DB 2;  Length 347;
  Best Local Similarity   100.0%;  Pred. No. 0.24;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       5 YPYDVPDYA 13
           |||||||||
Db      98 YPYDVPDYA 106


RESULT  13
S52176
hemagglutinin - influenza A virus (isolate A/England/247/93) (fragment)
C;Species: influenza A virus
A;Variety: isolate A/England/247/93
C;Date: 07-May-1995 #sequence_revision 23-Feb-1996 #text_change 23-Mar-2001
C;Accession: S52176
R;Ellis, J.S.
submitted to the EMBL Data Library, October 1994
A;Description: Antigenic and genetic variation in the haemagglutinin of recently circula
A;Reference number: S52173
A;Accession: S52176
A;Molecule type: mRNA
A;Residues: 1-347 <ELL>
A;Cross-references: EMBL:Z46394; NID:g609025; PIDN:CAA86529.1; PID:g609026
A;Experimental source: isolate A/England/247/93
C;Superfamily: influenza virus hemagglutinin


  Query Match             73.3%;  Score 55;  DB 2;  Length 347;
  Best Local Similarity   100.0%;  Pred. No. 0.24;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       5 YPYDVPDYA 13
           |||||||||
Db      98 YPYDVPDYA 106


RESULT  14
S52178
hemagglutinin - influenza A virus (isolate A/England/284/93) (fragment)
C;Species: influenza A virus
A;Variety: isolate A/England/284/93
C;Date: 07-May-1995 #sequence_revision 23-Feb-1996 #text_change 23-Mar-2001
C;Accession: S52178
R;Ellis, J.S.
submitted to the EMBL Data Library, October 1994
A;Description: Antigenic and genetic variation in the haemagglutinin of recently circ
A;Reference number: S52173
A;Accession: S52178
A;Molecule type: mRNA
A;Residues: 1-347 <ELL>
A;Cross-references: EMBL:Z46396; NID:g609029; PIDN:CAA86531.1; PID:g609030
A;Experimental source: isolate A/England/284/93
C;Superfamily: influenza virus hemagglutinin


  Query Match             73.3%;  Score 55;  DB 2;  Length 347;
  Best Local Similarity   100.0%;  Pred. No. 0.24;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       5 YPYDVPDYA 13
           |||||||||
Db      98 YPYDVPDYA 106


RESULT  15
S52179
hemagglutinin - influenza A virus (isolate A/England/286/93) (fragment)
C;Species: influenza A virus
A;Variety: isolate A/England/286/93
C;Date: 07-May-1995 #sequence_revision 23-Feb-1996 #text_change 23-Mar-2001
C;Accession: S52179
R;Ellis, J.S.
submitted to the EMBL Data Library, October 1994
A;Description: Antigenic and genetic variation in the haemagglutinin of recently circ
A;Reference number: S52173
A;Accession: S52179
A;Molecule type: mRNA
A;Residues: 1-347 <ELL>
A;Cross-references: EMBL:Z46397; NID:g609031; PIDN:CAA86532.1; PID:g609032
A;Experimental source: isolate A/England/286/93
C;Superfamily: influenza virus hemagglutinin


  Query Match             73.3%;  Score 55;  DB 2;  Length 347;
  Best Local Similarity   100.0%;  Pred. No. 0.24;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       5 YPYDVPDYA 13
           |||||||||
Db      98 YPYDVPDYA 106


Search completed: April 11, 2002, 10:16:16
Job time: 126 sec
```

GenCore version 4.5

OM protein - protein search, using sw model

Run on: April 11, 2002, 10:22:00 ; Search time 18.1 Seconds
(without alignments)
26.334 Million cell updates/sec

Title: US-09-284-787-3
Perfect score: 75
Sequence: 1 SGSGYPYDVPDYA 13

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 100059 seqs, 36664827 residues

Total number of hits satisfying chosen parameters: 100059

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	56	74.7	550	1	HEMA_IADHK	P43257 influenza a
2	56	74.7	550	1	HEMA_IAZH3	P11134 influenza a
3	56	74.7	565	1	HEMA_IAHK6	P19699 influenza a
4	55	73.3	249	1	HEMA_IAX31	P03438 influenza a
5	55	73.3	328	1	HEMA_IAEN6	P04664 influenza a
6	55	73.3	328	1	HEMA_IAQU7	P04663 influenza a
7	55	73.3	550	1	HEMA_IABAN	P03441 influenza a
8	55	73.3	550	1	HEMA_IADH1	P12582 influenza a
9	55	73.3	550	1	HEMA_IADH2	P12583 influenza a
10	55	73.3	550	1	HEMA_IADH3	P12584 influenza a
11	55	73.3	550	1	HEMA_IADH4	P12585 influenza a
12	55	73.3	550	1	HEMA_IADH6	P12587 influenza a
13	55	73.3	550	1	HEMA_IADH7	P12588 influenza a
14	55	73.3	550	1	HEMA_IADHL	P43258 influenza a
15	55	73.3	550	1	HEMA_IADHM	P43259 influenza a
16	55	73.3	550	1	HEMA_IAGHK	P43260 influenza a
17	55	73.3	550	1	HEMA_IAME6	P12589 influenza a
18	55	73.3	550	1	HEMA_IAZH2	P11133 influenza a
19	55	73.3	565	1	HEMA_IAHAL	P16994 influenza a
20	55	73.3	565	1	HEMA_IAHMI	P15658 influenza a
21	55	73.3	566	1	HEMA_IAAIC	P03437 influenza a
22	55	73.3	566	1	HEMA_IADM2	P26135 influenza a
23	55	73.3	566	1	HEMA_IAEN7	P03440 influenza a
24	55	73.3	566	1	HEMA_IAMAO	P26138 influenza a
25	55	73.3	566	1	HEMA_IAME1	P03449 influenza a
26	55	73.3	566	1	HEMA_IAME2	P03439 influenza a
27	55	73.3	566	1	HEMA_IANT6	P03436 influenza a
28	55	73.3	566	1	HEMA_IAUDO	P19106 influenza a
29	55	73.3	566	1	HEMA_IAZCO	P26139 influenza a
30	55	73.3	567	1	HEMA_IAVI7	P03435 influenza a
31	54	72.0	565	1	HEMA_IAHFO	P16995 influenza a
32	54	72.0	565	1	HEMA_IAHK7	P16996 influenza a
33	54	72.0	565	1	HEMA_IAHNM	P16997 influenza a
34	54	72.0	565	1	HEMA_IAHRO	P16998 influenza a
35	54	72.0	565	1	HEMA_IAHSA	P16999 influenza a
36	54	72.0	565	1	HEMA_IAHSU	Q08011 influenza a
37	54	72.0	565	1	HEMA_IAHTE	P17001 influenza a
38	54	72.0	565	1	HEMA_IAHUR	P17002 influenza a
39	54	72.0	566	1	HEMA_IADU3	P03442 influenza a
40	51	68.0	566	1	HEMA_IADA3	P26134 influenza a
41	51	68.0	566	1	HEMA_IAZUK	P26141 influenza a
42	50	66.7	565	1	HEMA_IAHTO	P17000 influenza a
43	47	62.7	550	1	HEMA_IADH5	P12586 influenza a
44	47	62.7	563	1	HEMA_IAMAA	P26137 influenza a
45	47	62.7	564	1	HEMA_IABUD	P19694 influenza a

ALIGNMENTS

RESULT 1
HEMA_IADHK
ID HEMA_IADHK STANDARD; PRT; 550 AA.
AC P43257;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE HEMAGGLUTININ PRECURSOR [CONTAINS: HEMAGGLUTININ HA1 CHAIN;
DE HEMAGGLUTININ HA2 CHAIN] (FRAGMENT).
GN HA.
OS Influenza A virus (strain A/Duck/Hong Kong/7/75).
OC Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;
OC Influenza virus A and B group; Influenza A viruses;
OC Influenza A virus.
OX NCBI_TaxID=11364;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91341491; PubMed=1875195;
RA Yasuda J., Shortridge K.F., Shimizu Y., Kida H.;
RT "Molecular evidence for a role of domestic ducks in the introduction
RT of avian H3 influenza viruses to pigs in southern China, where the
RT A/Hong Kong/68 (H3N2) strain emerged.";
RL J. Gen. Virol. 72:2007-2010(1991).
CC -!- FUNCTION: HEMAGGLUTININ IS RESPONSIBLE FOR ATTACHING THE VIRUS TO
CC CELL RECEPTORS AND FOR INITIATING INFECTION.
CC -!- SUBUNIT: HOMOTRIMER. EACH OF THE MONOMERS IS FORMED BY TWO CHAINS
CC (HA1 AND HA2) LINKED BY A DISULFIDE BOND.
CC -!- SIMILARITY: BELONGS TO THE INFLUENZA HEMAGGLUTININ FAMILY.
CC ---

CC ---
DR EMBL; D00929; BAA00769.1; -.
DR HSSP; P03437; 5HMG.
DR InterPro; IPR001364; Hemagglutn.
DR Pfam; PF00509; Hemagglutinin; 1.
DR ProDom; PD000225; Hemagglutn; 1.
KW Envelope protein; Hemagglutinin; Glycoprotein.
FT NON_TER 1 1
FT CHAIN 1 328 HEMAGGLUTININ HA1 CHAIN.
FT CHAIN 330 550 HEMAGGLUTININ HA2 CHAIN.
FT CARBOHYD 8 8 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 22 22 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 38 38 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 165 165 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 285 285 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 483 483 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 550 AA; 61549 MW; 864639B829FE1BA9 CRC64;

Query Match 74.7%; Score 56; DB 1; Length 550;

```
  Best Local Similarity   90.9%;  Pred. No. 0.084;
  Matches   10;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       3 SGYPYDVPDYA 13
           | ||||||||||
Db      96 SCYPYDVPDYA 106


RESULT   2
HEMA_IAZH3
ID   HEMA_IAZH3     STANDARD;      PRT;   550 AA.
AC   P11134; Q84025; Q84026;
DT   01-JUL-1989 (Rel. 11, Created)
DT   01-JUL-1989 (Rel. 11, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   HEMAGGLUTININ [CONTAINS: HEMAGGLUTININ HA1 CHAIN; HEMAGGLUTININ HA2
DE   CHAIN] (FRAGMENT).
GN   HA.
OS   Influenza A virus (strain A/Swine/Hong Kong/126/82).
OC   Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;
OC   Influenza virus A and B group; Influenza A viruses;
OC   Influenza A virus.
OX   NCBI_TaxID=11498;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=88101364; PubMed=3336940;
RA   Kida H., Shortridge K.F., Webster R.G.;
RT   "Origin of the hemagglutinin gene of H3N2 influenza viruses from pigs
RT   in China.";
RL   Virology 162:160-166(1988).
CC   -!- FUNCTION: HEMAGGLUTININ IS RESPONSIBLE FOR ATTACHING THE VIRUS TO
CC       CELL RECEPTORS AND FOR INITIATING INFECTION.
CC   -!- SUBUNIT: HOMOTRIMER. EACH OF THE MONOMERS IS FORMED BY TWO CHAINS
CC       (HA1 AND HA2) LINKED BY A DISULFIDE BOND.
CC   -!- SIMILARITY: BELONGS TO THE INFLUENZA HEMAGGLUTININ FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M19056; AAA43211.1; ALT_TERM.
DR   PIR; A29971; HMIVS2.
DR   HSSP; P03437; 2HMG.
DR   InterPro; IPR001364; Hemagglutn.
DR   Pfam; PF00509; Hemagglutinin; 1.
DR   ProDom; PD000225; Hemagglutn; 1.
KW   Hemagglutinin; Envelope protein; Glycoprotein.
FT   NON_TER       1      1
FT   CHAIN         1    328       HEMAGGLUTININ HA1 CHAIN.
FT   CHAIN       330    550       HEMAGGLUTININ HA2 CHAIN.
FT   CARBOHYD      8      8       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     22     22       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     38     38       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    165    165       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    285    285       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    483    483       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   550 AA;  61580 MW;  991F6D8BC02F24F2 CRC64;


  Query Match             74.7%;  Score 56;  DB 1;  Length 550;
  Best Local Similarity   90.9%;  Pred. No. 0.084;
  Matches   10;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       3 SGYPYDVPDYA 13
           | ||||||||||
Db      96 SCYPYDVPDYA 106


RESULT   3
HEMA_IAHK6
ID   HEMA_IAHK6     STANDARD;      PRT;   565 AA.
AC   P19699;
DT   01-FEB-1991 (Rel. 17, Created)
DT   01-FEB-1991 (Rel. 17, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   HEMAGGLUTININ PRECURSOR [CONTAINS: HEMAGGLUTININ HA1 CHAIN;
DE   HEMAGGLUTININ HA2 CHAIN].
GN   HA.
OS   Influenza A virus (strain A/Equine/Kentucky/2/86).
OC   Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;
OC   Influenza virus A and B group; Influenza A viruses;
OC   Influenza A virus.
OX   NCBI_TaxID=11403;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=89204899; PubMed=2705299;
RA   Kawaoka Y., Bean W.J., Webster R.G.;
RT   "Evolution of the hemagglutinin of equine H3 influenza viruses.";
RL   Virology 169:283-292(1989).
CC   -!- FUNCTION: HEMAGGLUTININ IS RESPONSIBLE FOR ATTACHING THE VIRUS TO
CC       CELL RECEPTORS AND FOR INITIATING INFECTION.
CC   -!- SUBUNIT: HOMOTRIMER. EACH OF THE MONOMERS IS FORMED BY TWO CHAINS
CC       (HA1 AND HA2) LINKED BY A DISULFIDE BOND.
CC   -!- SIMILARITY: BELONGS TO THE INFLUENZA HEMAGGLUTININ FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M24727; -; NOT_ANNOTATED_CDS.
DR   PIR; A34065; HMIVET.
DR   HSSP; P03437; 1HTM.
DR   InterPro; IPR001364; Hemagglutn.
DR   Pfam; PF00509; Hemagglutinin; 1.
DR   PRINTS; PR00329; HEMAGGLUTN12.
DR   PRINTS; PR00330; HEMAGGLUTN1.
DR   PRINTS; PR00331; HEMAGGLUTN2.
DR   ProDom; PD000225; Hemagglutn; 1.
KW   Envelope protein; Hemagglutinin; Glycoprotein; Signal.
FT   SIGNAL        1     16
FT   CHAIN        17    343       HEMAGGLUTININ HA1 CHAIN.
FT   CHAIN       345    565       HEMAGGLUTININ HA2 CHAIN.
FT   CARBOHYD     23     23       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     37     37       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     53     53       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     68     68       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     78     78       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    180    180       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    300    300       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    498    498       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   565 AA;  63610 MW;  2038CC1C6C9B88C5 CRC64;


  Query Match             74.7%;  Score 56;  DB 1;  Length 565;
  Best Local Similarity   69.2%;  Pred. No. 0.086;
  Matches    9;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 SGSGYPYDVPDYA 13
           | : ||||:||||
Db     109 SSNCYPYDIPDYA 121


RESULT   4
HEMA_IAX31
ID   HEMA_IAX31     STANDARD;      PRT;   249 AA.
AC   P03438;
DT   21-JUL-1986 (Rel. 01, Created)
DT   21-JUL-1986 (Rel. 01, Last sequence update)
```

```
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   HEMAGGLUTININ PRECURSOR [CONTAINS: HEMAGGLUTININ HA1 CHAIN]
DE   (FRAGMENT).
GN   HA.
OS   Influenza A virus (strain X-31).
OC   Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;
OC   Influenza virus A and B group; Influenza A viruses;
OC   Influenza A virus.
OX   NCBI_TaxID=11489;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=81030852; PubMed=7421990;
RA   Gething M.-J., Bye J., Skehel J.J., Waterfield M.;
RT   "Cloning and DNA sequence of double-stranded copies of haemagglutinin
RT   genes from H2 and H3 strains elucidates antigenic shift and drift in
RT   human influenza virus.";
RL   Nature 287:301-306(1980).
CC   -!- FUNCTION: HEMAGGLUTININ IS RESPONSIBLE FOR ATTACHING THE VIRUS TO
CC       CELL RECEPTORS AND FOR INITIATING INFECTION.
CC   -!- SUBUNIT: HOMOTRIMER. EACH OF THE MONOMERS IS FORMED BY TWO CHAINS
CC       (HA1 AND HA2) LINKED BY A DISULFIDE BOND.
CC   -!- SIMILARITY: BELONGS TO THE INFLUENZA HEMAGGLUTININ FAMILY.
DR   PIR; A04051; HMIVH.
DR   PIR; A93233; A93233.
DR   HSSP; P03437; 5HMG.
DR   InterPro; IPR001364; Hemagglutn.
DR   Pfam; PF00509; Hemagglutinin; 1.
DR   ProDom; PD000225; Hemagglutn; 1.
KW   Envelope protein; Hemagglutinin; Glycoprotein; Signal.
FT   SIGNAL        1     16
FT   CHAIN        17   >249       HEMAGGLUTININ HA1 CHAIN.
FT   CARBOHYD     24     24       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     38     38       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     54     54       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     97     97       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    181    181       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   NON_TER     249    249
SQ   SEQUENCE   249 AA;  27373 MW;  9537AA970BB79183 CRC64;

  Query Match             73.3%;  Score 55;  DB 1;  Length 249;
  Best Local Similarity   100.0%;  Pred. No. 0.055;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       5 YPYDVPDYA 13
           |||||||||
Db     114 YPYDVPDYA 122


RESULT   5
HEMA_IAEN6
ID   HEMA_IAEN6     STANDARD;      PRT;   328 AA.
AC   P04664;
DT   13-AUG-1987 (Rel. 05, Created)
DT   13-AUG-1987 (Rel. 05, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   HEMAGGLUTININ [CONTAINS: HEMAGGLUTININ HA1 CHAIN] (FRAGMENT).
GN   HA.
OS   Influenza A virus (strain A/England/878/69).
OC   Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;
OC   Influenza virus A and B group; Influenza A viruses;
OC   Influenza A virus.
OX   NCBI_TaxID=11377;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=81194918; PubMed=6164798;
RA   Sleigh M.J., Both G.W., Underwood P.A., Bender V.J.;
RT   "Antigenic drift in the hemagglutinin of the Hong Kong influenza
RT   subtype: correlation of amino acid changes with alterations in viral
RT   antigenicity.";
RL   J. Virol. 37:845-853(1981).
CC   -!- FUNCTION: HEMAGGLUTININ IS RESPONSIBLE FOR ATTACHING THE VIRUS TO
CC       CELL RECEPTORS AND FOR INITIATING INFECTION.
CC   -!- SUBUNIT: HOMOTRIMER. EACH OF THE MONOMERS IS FORMED BY TWO CHAINS
CC       (HA1 AND HA2) LINKED BY A DISULFIDE BOND.
CC   -!- SIMILARITY: BELONGS TO THE INFLUENZA HEMAGGLUTININ FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; K03335; AAA43184.1; -.
DR   InterPro; IPR001364; Hemagglutn.
DR   Pfam; PF00509; Hemagglutinin; 1.
DR   ProDom; PD000225; Hemagglutn; 1.
KW   Envelope protein; Hemagglutinin; Glycoprotein.
FT   NON_TER       1      1
FT   CHAIN         1    328       HEMAGGLUTININ HA1 CHAIN.
FT   CARBOHYD      8      8       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     22     22       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     38     38       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     63     63       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    165    165       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    285    285       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   NON_TER     328    328
SQ   SEQUENCE   328 AA;  36072 MW;  9C3A86B3A8D856FE CRC64;


  Query Match             73.3%;  Score 55;  DB 1;  Length 328;
  Best Local Similarity   100.0%;  Pred. No. 0.072;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       5 YPYDVPDYA 13
           |||||||||
Db      98 YPYDVPDYA 106


RESULT   6
HEMA_IAQU7
ID   HEMA_IAQU7     STANDARD;      PRT;   328 AA.
AC   P04663;
DT   13-AUG-1987 (Rel. 05, Created)
DT   13-AUG-1987 (Rel. 05, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   HEMAGGLUTININ [CONTAINS: HEMAGGLUTININ HA1 CHAIN] (FRAGMENT).
GN   HA.
OS   Influenza A virus (strain A/Qu/7/70).
OC   Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;
OC   Influenza virus A and B group; Influenza A viruses;
OC   Influenza A virus.
OX   NCBI_TaxID=11456;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=81194918; PubMed=6164798;
RA   Sleigh M.J., Both G.W., Underwood P.A., Bender V.J.;
RT   "Antigenic drift in the hemagglutinin of the Hong Kong influenza
RT   subtype: correlation of amino acid changes with alterations in viral
RT   antigenicity.";
RL   J. Virol. 37:845-853(1981).
CC   -!- FUNCTION: HEMAGGLUTININ IS RESPONSIBLE FOR ATTACHING THE VIRUS TO
CC       CELL RECEPTORS AND FOR INITIATING INFECTION.
CC   -!- SUBUNIT: HOMOTRIMER. EACH OF THE MONOMERS IS FORMED BY TWO CHAINS
CC       (HA1 AND HA2) LINKED BY A DISULFIDE BOND.
CC   -!- SIMILARITY: BELONGS TO THE INFLUENZA HEMAGGLUTININ FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
```

```
CC   ----------------------------------------------------------------------
DR   EMBL; K03338; AAA43195.1; -.
DR   InterPro; IPR001364; Hemagglutn.
DR   Pfam; PF00509; Hemagglutinin; 1.
DR   ProDom; PD000225; Hemagglutn; 1.
KW   Envelope protein; Hemagglutinin; Glycoprotein.
FT   NON_TER       1      1
FT   CHAIN         1    328       HEMAGGLUTININ HA1 CHAIN.
FT   CARBOHYD      8      8       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     22     22       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     38     38       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     81     81       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    165    165       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    285    285       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   NON_TER     328    328
SQ   SEQUENCE   328 AA;  36059 MW;  E30A962FD6DD805B CRC64;


  Query Match             73.3%;  Score 55;  DB 1;  Length 328;
  Best Local Similarity   100.0%;  Pred. No. 0.072;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       5 YPYDVPDYA 13
           |||||||||
Db      98 YPYDVPDYA 106


RESULT   7
HEMA_IABAN
ID   HEMA_IABAN     STANDARD;      PRT;   550 AA.
AC   P03441; Q83961; Q83962;
DT   21-JUL-1986 (Rel. 01, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   HEMAGGLUTININ PRECURSOR [CONTAINS: HEMAGGLUTININ HA1 CHAIN;
DE   HEMAGGLUTININ HA2 CHAIN] (FRAGMENT).
GN   HA.
OS   Influenza A virus (strain A/Bangkok/1/79).
OC   Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;
OC   Influenza virus A and B group; Influenza A viruses;
OC   Influenza A virus.
OX   NCBI_TaxID=11325;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=82033259; PubMed=6169840;
RA   Both G.W., Sleigh M.J.;
RT   "Conservation and variation in the hemagglutinins of Hong Kong
RT   subtype influenza viruses during antigenic drift.";
RL   J. Virol. 39:663-672(1981).
CC   -!- FUNCTION: HEMAGGLUTININ IS RESPONSIBLE FOR ATTACHING THE VIRUS TO
CC       CELL RECEPTORS AND FOR INITIATING INFECTION.
CC   -!- SUBUNIT: HOMOTRIMER. EACH OF THE MONOMERS IS FORMED BY TWO CHAINS
CC       (HA1 AND HA2) LINKED BY A DISULFIDE BOND.
CC   -!- SIMILARITY: BELONGS TO THE INFLUENZA HEMAGGLUTININ FAMILY.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; J02092; AAA43182.1; -.
DR   PIR; A04052; HMIV6.
DR   InterPro; IPR001364; Hemagglutn.
DR   Pfam; PF00509; Hemagglutinin; 1.
DR   ProDom; PD000225; Hemagglutn; 1.
KW   Envelope protein; Hemagglutinin; Glycoprotein.
FT   NON_TER       1      1
FT   CHAIN         1    328       HEMAGGLUTININ HA1 CHAIN.
FT   CHAIN       330    550       HEMAGGLUTININ HA2 CHAIN.
FT   CARBOHYD      8      8       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     22     22       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     38     38       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     63     63       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    126    126       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    165    165       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    285    285       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    483    483       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   550 AA;  61699 MW;  FBD2EC200689CBE5 CRC64;


  Query Match             73.3%;  Score 55;  DB 1;  Length 550;
  Best Local Similarity   100.0%;  Pred. No. 0.12;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       5 YPYDVPDYA 13
           |||||||||
Db      98 YPYDVPDYA 106


RESULT   8
HEMA_IADH1
ID   HEMA_IADH1     STANDARD;      PRT;   550 AA.
AC   P12582; Q84021; Q84022;
DT   01-OCT-1989 (Rel. 12, Created)
DT   01-OCT-1989 (Rel. 12, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   HEMAGGLUTININ PRECURSOR [CONTAINS: HEMAGGLUTININ HA1 CHAIN;
DE   HEMAGGLUTININ HA2 CHAIN] (FRAGMENT).
GN   HA.
OS   Influenza A virus (strain A/Duck/Hokkaido/5/77).
OC   Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;
OC   Influenza virus A and B group; Influenza A viruses;
OC   Influenza A virus.
OX   NCBI_TaxID=11357;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=87265458; PubMed=2440178;
RA   Kida H., Kawaoka Y., Naeve C.W., Webster R.G.;
RT   "Antigenic and genetic conservation of H3 influenza virus in wild
RT   ducks.";
RL   Virology 159:109-119(1987).
CC   -!- FUNCTION: HEMAGGLUTININ IS RESPONSIBLE FOR ATTACHING THE VIRUS TO
CC       CELL RECEPTORS AND FOR INITIATING INFECTION.
CC   -!- SUBUNIT: HOMOTRIMER. EACH OF THE MONOMERS IS FORMED BY TWO CHAINS
CC       (HA1 AND HA2) LINKED BY A DISULFIDE BOND.
CC   -!- SIMILARITY: BELONGS TO THE INFLUENZA HEMAGGLUTININ FAMILY.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; M16737; AAA43143.1; -.
DR   PIR; A27813; HMIV77.
DR   InterPro; IPR001364; Hemagglutn.
DR   Pfam; PF00509; Hemagglutinin; 1.
DR   ProDom; PD000225; Hemagglutn; 1.
KW   Envelope protein; Hemagglutinin; Glycoprotein.
FT   NON_TER       1      1
FT   CHAIN         1    328       HEMAGGLUTININ HA1 CHAIN.
FT   CHAIN       330    550       HEMAGGLUTININ HA2 CHAIN.
FT   CARBOHYD      8      8       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     22     22       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     38     38       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    165    165       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    285    285       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    483    483       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   550 AA;  61705 MW;  7E7ACFE716FC969A CRC64;
```

```
  Query Match             73.3%;  Score 55;  DB 1;  Length 550;
  Best Local Similarity   100.0%;  Pred. No. 0.12;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      5 YPYDVPDYA 13
          |||||||||
Db     98 YPYDVPDYA 106


RESULT  9
HEMA_IADH2
ID   HEMA_IADH2     STANDARD;      PRT;   550 AA.
AC   P12583; Q84011;
DT   01-OCT-1989 (Rel. 12, Created)
DT   01-APR-1990 (Rel. 14, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   HEMAGGLUTININ PRECURSOR [CONTAINS: HEMAGGLUTININ HA1 CHAIN;
DE   HEMAGGLUTININ HA2 CHAIN] (FRAGMENT).
GN   HA.
OS   Influenza A virus (strain A/Duck/Hokkaido/8/80).
OC   Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;
OC   Influenza virus A and B group; Influenza A viruses;
OC   Influenza A virus.
OX   NCBI_TaxID=11358;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=87265458; PubMed=2440178;
RA   Kida H., Kawaoka Y., Naeve C.W., Webster R.G.;
RT   "Antigenic and genetic conservation of H3 influenza virus in wild
RT   ducks.";
RL   Virology 159:109-119(1987).
CC   -!- FUNCTION: HEMAGGLUTININ IS RESPONSIBLE FOR ATTACHING THE VIRUS TO
CC       CELL RECEPTORS AND FOR INITIATING INFECTION.
CC   -!- SUBUNIT: HOMOTRIMER. EACH OF THE MONOMERS IS FORMED BY TWO CHAINS
CC       (HA1 AND HA2) LINKED BY A DISULFIDE BOND.
CC   -!- SIMILARITY: BELONGS TO THE INFLUENZA HEMAGGLUTININ FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M16738; AAA43144.1; -.
DR   PIR; B27813; HMIV80.
DR   InterPro; IPR001364; Hemagglutn.
DR   Pfam; PF00509; Hemagglutinin; 1.
DR   ProDom; PD000225; Hemagglutn; 1.
KW   Envelope protein; Hemagglutinin; Glycoprotein.
FT   NON_TER       1      1
FT   CHAIN         1    328       HEMAGGLUTININ HA1 CHAIN.
FT   CHAIN       330    550       HEMAGGLUTININ HA2 CHAIN.
FT   CARBOHYD      8      8       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     22     22       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     38     38       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    165    165       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    285    285       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    483    483       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CONFLICT    137    137       K -> N (IN PIR DATA BANK).
SQ   SEQUENCE   550 AA;  61659 MW;  A107023ACC9CC353 CRC64;


  Query Match             73.3%;  Score 55;  DB 1;  Length 550;
  Best Local Similarity   100.0%;  Pred. No. 0.12;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      5 YPYDVPDYA 13
          |||||||||
Db     98 YPYDVPDYA 106


RESULT  10
HEMA_IADH3
ID   HEMA_IADH3     STANDARD;      PRT;   550 AA.
AC   P12584; Q84012; Q89793;
DT   01-OCT-1989 (Rel. 12, Created)
DT   01-OCT-1989 (Rel. 12, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   HEMAGGLUTININ PRECURSOR [CONTAINS: HEMAGGLUTININ HA1 CHAIN;
DE   HEMAGGLUTININ HA2 CHAIN] (FRAGMENT).
GN   HA.
OS   Influenza A virus (strain A/Duck/Hokkaido/33/80).
OC   Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;
OC   Influenza virus A and B group; Influenza A viruses;
OC   Influenza A virus.
OX   NCBI_TaxID=11359;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=87265458; PubMed=2440178;
RA   Kida H., Kawaoka Y., Naeve C.W., Webster R.G.;
RT   "Antigenic and genetic conservation of H3 influenza virus in wild
RT   ducks.";
RL   Virology 159:109-119(1987).
CC   -!- FUNCTION: HEMAGGLUTININ IS RESPONSIBLE FOR ATTACHING THE VIRUS TO
CC       CELL RECEPTORS AND FOR INITIATING INFECTION.
CC   -!- SUBUNIT: HOMOTRIMER. EACH OF THE MONOMERS IS FORMED BY TWO CHAINS
CC       (HA1 AND HA2) LINKED BY A DISULFIDE BOND.
CC   -!- SIMILARITY: BELONGS TO THE INFLUENZA HEMAGGLUTININ FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M16739; AAA43145.1; -.
DR   PIR; C27813; HMIV33.
DR   InterPro; IPR001364; Hemagglutn.
DR   Pfam; PF00509; Hemagglutinin; 1.
DR   ProDom; PD000225; Hemagglutn; 1.
KW   Envelope protein; Hemagglutinin; Glycoprotein.
FT   NON_TER       1      1
FT   CHAIN         1    328       HEMAGGLUTININ HA1 CHAIN.
FT   CHAIN       330    550       HEMAGGLUTININ HA2 CHAIN.
FT   CARBOHYD      8      8       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     22     22       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     38     38       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    165    165       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    285    285       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    483    483       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   550 AA;  61577 MW;  6C30BF67CFDCB7DE CRC64;


  Query Match             73.3%;  Score 55;  DB 1;  Length 550;
  Best Local Similarity   100.0%;  Pred. No. 0.12;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      5 YPYDVPDYA 13
          |||||||||
Db     98 YPYDVPDYA 106


RESULT  11
HEMA_IADH4
ID   HEMA_IADH4     STANDARD;      PRT;   550 AA.
AC   P12585; Q84013; Q84014;
DT   01-OCT-1989 (Rel. 12, Created)
DT   01-OCT-1989 (Rel. 12, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   HEMAGGLUTININ PRECURSOR [CONTAINS: HEMAGGLUTININ HA1 CHAIN;
DE   HEMAGGLUTININ HA2 CHAIN] (FRAGMENT).
GN   HA.
```

```
OS   Influenza A virus (strain A/Duck/Hokkaido/7/82).
OC   Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;
OC   Influenza virus A and B group; Influenza A viruses;
OC   Influenza A virus.
OX   NCBI_TaxID=11360;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=87265458; PubMed=2440178;
RA   Kida H., Kawaoka Y., Naeve C.W., Webster R.G.;
RT   "Antigenic and genetic conservation of H3 influenza virus in wild
RT   ducks.";
RL   Virology 159:109-119(1987).
CC   -!- FUNCTION: HEMAGGLUTININ IS RESPONSIBLE FOR ATTACHING THE VIRUS TO
CC       CELL RECEPTORS AND FOR INITIATING INFECTION.
CC   -!- SUBUNIT: HOMOTRIMER. EACH OF THE MONOMERS IS FORMED BY TWO CHAINS
CC       (HA1 AND HA2) LINKED BY A DISULFIDE BOND.
CC   -!- SIMILARITY: BELONGS TO THE INFLUENZA HEMAGGLUTININ FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M16740; AAA43146.1; -.
DR   PIR; D27813; HMIV89.
DR   InterPro; IPR001364; Hemagglutn.
DR   Pfam; PF00509; Hemagglutinin; 1.
DR   ProDom; PD000225; Hemagglutn; 1.
KW   Envelope protein; Hemagglutinin; Glycoprotein.
FT   NON_TER       1      1
FT   CHAIN         1    328       HEMAGGLUTININ HA1 CHAIN.
FT   CHAIN       330    550       HEMAGGLUTININ HA2 CHAIN.
FT   CARBOHYD      8      8       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     22     22       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     38     38       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    165    165       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    285    285       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    483    483       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   550 AA;  61664 MW;  A16B2CF8CBBBD9D0 CRC64;


  Query Match             73.3%;  Score 55;  DB 1;  Length 550;
  Best Local Similarity   100.0%;  Pred. No. 0.12;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       5 YPYDVPDYA 13
           |||||||||
Db      98 YPYDVPDYA 106


RESULT  12
HEMA_IADH6
ID   HEMA_IADH6     STANDARD;      PRT;   550 AA.
AC   P12587; Q84017;
DT   01-OCT-1989 (Rel. 12, Created)
DT   01-APR-1990 (Rel. 14, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   HEMAGGLUTININ PRECURSOR [CONTAINS: HEMAGGLUTININ HA1 CHAIN;
DE   HEMAGGLUTININ HA2 CHAIN] (FRAGMENT).
GN   HA.
OS   Influenza A virus (strain A/Duck/Hokkaido/9/85).
OC   Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;
OC   Influenza virus A and B group; Influenza A viruses;
OC   Influenza A virus.
OX   NCBI_TaxID=11362;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=87265458; PubMed=2440178;
RA   Kida H., Kawaoka Y., Naeve C.W., Webster R.G.;
RT   "Antigenic and genetic conservation of H3 influenza virus in wild
RT   ducks.";
RL   Virology 159:109-119(1987).
CC   -!- FUNCTION: HEMAGGLUTININ IS RESPONSIBLE FOR ATTACHING THE VIRUS TO
CC       CELL RECEPTORS AND FOR INITIATING INFECTION.
CC   -!- SUBUNIT: HOMOTRIMER. EACH OF THE MONOMERS IS FORMED BY TWO CHAINS
CC       (HA1 AND HA2) LINKED BY A DISULFIDE BOND.
CC   -!- SIMILARITY: BELONGS TO THE INFLUENZA HEMAGGLUTININ FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M16742; AAA43148.1; -.
DR   PIR; F27813; HMIV98.
DR   InterPro; IPR001364; Hemagglutn.
DR   Pfam; PF00509; Hemagglutinin; 1.
DR   ProDom; PD000225; Hemagglutn; 1.
KW   Envelope protein; Hemagglutinin; Glycoprotein.
FT   NON_TER       1      1
FT   CHAIN         1    328       HEMAGGLUTININ HA1 CHAIN.
FT   CHAIN       330    550       HEMAGGLUTININ HA2 CHAIN.
FT   CARBOHYD     22     22       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     38     38       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    165    165       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    285    285       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    483    483       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CONFLICT      8      8       Y -> N (IN PIR DATA BANK).
SQ   SEQUENCE   550 AA;  61711 MW;  67BCD85F44736CFE CRC64;


  Query Match             73.3%;  Score 55;  DB 1;  Length 550;
  Best Local Similarity   100.0%;  Pred. No. 0.12;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       5 YPYDVPDYA 13
           |||||||||
Db      98 YPYDVPDYA 106


RESULT  13
HEMA_IADH7
ID   HEMA_IADH7     STANDARD;      PRT;   550 AA.
AC   P12588; Q84018; Q89470;
DT   01-OCT-1989 (Rel. 12, Created)
DT   01-OCT-1989 (Rel. 12, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   HEMAGGLUTININ PRECURSOR [CONTAINS: HEMAGGLUTININ HA1 CHAIN;
DE   HEMAGGLUTININ HA2 CHAIN] (FRAGMENT).
GN   HA.
OS   Influenza A virus (strain A/Duck/Hokkaido/10/85).
OC   Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;
OC   Influenza virus A and B group; Influenza A viruses;
OC   Influenza A virus.
OX   NCBI_TaxID=11363;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=87265458; PubMed=2440178;
RA   Kida H., Kawaoka Y., Naeve C.W., Webster R.G.;
RT   "Antigenic and genetic conservation of H3 influenza virus in wild
RT   ducks.";
RL   Virology 159:109-119(1987).
CC   -!- FUNCTION: HEMAGGLUTININ IS RESPONSIBLE FOR ATTACHING THE VIRUS TO
CC       CELL RECEPTORS AND FOR INITIATING INFECTION.
CC   -!- SUBUNIT: HOMOTRIMER. EACH OF THE MONOMERS IS FORMED BY TWO CHAINS
CC       (HA1 AND HA2) LINKED BY A DISULFIDE BOND.
CC   -!- SIMILARITY: BELONGS TO THE INFLUENZA HEMAGGLUTININ FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
```

```
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; M16743; AAA43149.1; -.
DR   PIR; G27813; HMIV15.
DR   InterPro; IPR001364; Hemagglutn.
DR   Pfam; PF00509; Hemagglutinin; 1.
DR   ProDom; PD000225; Hemagglutn; 1.
KW   Envelope protein; Hemagglutinin; Glycoprotein.
FT   NON_TER       1      1
FT   CHAIN         1    328       HEMAGGLUTININ HA1 CHAIN.
FT   CHAIN       330    550       HEMAGGLUTININ HA2 CHAIN.
FT   CARBOHYD      8      8       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     22     22       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     38     38       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    165    165       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    285    285       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    483    483       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   550 AA;  61761 MW;  6EF81793281D53EB CRC64;


  Query Match             73.3%;  Score 55;  DB 1;  Length 550;
  Best Local Similarity   100.0%;  Pred. No. 0.12;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      5 YPYDVPDYA 13
          |||||||||
Db     98 YPYDVPDYA 106


RESULT  14
HEMA_IADHL
ID   HEMA_IADHL     STANDARD;      PRT;   550 AA.
AC   P43258;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   HEMAGGLUTININ PRECURSOR [CONTAINS: HEMAGGLUTININ HA1 CHAIN;
DE   HEMAGGLUTININ HA2 CHAIN] (FRAGMENT).
GN   HA.
OS   Influenza A virus (strain A/Duck/Hong Kong/64/76).
OC   Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;
OC   Influenza virus A and B group; Influenza A viruses;
OC   Influenza A virus.
OX   NCBI_TaxID=45412;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=91341491; PubMed=1875195;
RA   Yasuda J., Shortridge K.F., Shimizu Y., Kida H.;
RT   "Molecular evidence for a role of domestic ducks in the introduction
RT   of avian H3 influenza viruses to pigs in southern China, where the
RT   A/Hong Kong/68 (H3N2) strain emerged.";
RL   J. Gen. Virol. 72:2007-2010(1991).
CC   -!- FUNCTION: HEMAGGLUTININ IS RESPONSIBLE FOR ATTACHING THE VIRUS TO
CC       CELL RECEPTORS AND FOR INITIATING INFECTION.
CC   -!- SUBUNIT: HOMOTRIMER. EACH OF THE MONOMERS IS FORMED BY TWO CHAINS
CC       (HA1 AND HA2) LINKED BY A DISULFIDE BOND.
CC   -!- SIMILARITY: BELONGS TO THE INFLUENZA HEMAGGLUTININ FAMILY.
CC   ---------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; D00931; BAA00771.1; -.
DR   HSSP; P03437; 2HMG.
DR   InterPro; IPR001364; Hemagglutn.
DR   Pfam; PF00509; Hemagglutinin; 1.
DR   ProDom; PD000225; Hemagglutn; 1.
KW   Envelope protein; Hemagglutinin; Glycoprotein.
FT   NON_TER       1      1
FT   CHAIN         1    328       HEMAGGLUTININ HA1 CHAIN.
FT   CHAIN       330    550       HEMAGGLUTININ HA2 CHAIN.
FT   CARBOHYD      8      8       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     22     22       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     38     38       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    165    165       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    285    285       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    483    483       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   550 AA;  61718 MW;  A351C56789E4BE9A CRC64;


  Query Match             73.3%;  Score 55;  DB 1;  Length 550;
  Best Local Similarity   100.0%;  Pred. No. 0.12;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      5 YPYDVPDYA 13
          |||||||||
Db     98 YPYDVPDYA 106


RESULT  15
HEMA_IADHM
ID   HEMA_IADHM     STANDARD;      PRT;   550 AA.
AC   P43259;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   HEMAGGLUTININ PRECURSOR [CONTAINS: HEMAGGLUTININ HA1 CHAIN;
DE   HEMAGGLUTININ HA2 CHAIN] (FRAGMENT).
GN   HA.
OS   Influenza A virus (strain A/Duck/Hong Kong/231/77).
OC   Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;
OC   Influenza virus A and B group; Influenza A viruses;
OC   Influenza A virus.
OX   NCBI_TaxID=45411;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=91341491; PubMed=1875195;
RA   Yasuda J., Shortridge K.F., Shimizu Y., Kida H.;
RT   "Molecular evidence for a role of domestic ducks in the introduction
RT   of avian H3 influenza viruses to pigs in southern China, where the
RT   A/Hong Kong/68 (H3N2) strain emerged.";
RL   J. Gen. Virol. 72:2007-2010(1991).
CC   -!- FUNCTION: HEMAGGLUTININ IS RESPONSIBLE FOR ATTACHING THE VIRUS TO
CC       CELL RECEPTORS AND FOR INITIATING INFECTION.
CC   -!- SUBUNIT: HOMOTRIMER. EACH OF THE MONOMERS IS FORMED BY TWO CHAINS
CC       (HA1 AND HA2) LINKED BY A DISULFIDE BOND.
CC   -!- SIMILARITY: BELONGS TO THE INFLUENZA HEMAGGLUTININ FAMILY.
CC   ---------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; D00932; BAA00772.1; -.
DR   InterPro; IPR001364; Hemagglutn.
DR   Pfam; PF00509; Hemagglutinin; 1.
DR   ProDom; PD000225; Hemagglutn; 1.
KW   Envelope protein; Hemagglutinin; Glycoprotein.
FT   NON_TER       1      1
FT   CHAIN         1    328       HEMAGGLUTININ HA1 CHAIN.
FT   CHAIN       330    550       HEMAGGLUTININ HA2 CHAIN.
FT   CARBOHYD     22     22       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     38     38       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    165    165       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    285    285       N-LINKED (GLCNAC. . .) (POTENTIAL).
```

```
FT   CARBOHYD    483    483       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   550 AA;  61762 MW;  6FEF8B49488C191A CRC64;


  Query Match             73.3%;  Score 55;  DB 1;  Length 550;
  Best Local Similarity   100.0%;  Pred. No. 0.12;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       5 YPYDVPDYA 13
           |||||||||
Db      98 YPYDVPDYA 106


Search completed: April 11, 2002, 10:22:01
Job time: 431 sec
```

GenCore version 4.5

OM protein - protein search, using sw model

Run on: April 11, 2002, 10:22:57 ; Search time 50.04 Seconds
(without alignments)
38.000 Million cell updates/sec

Title: US-09-284-787-3
Perfect score: 75
Sequence: 1 SGSGYPYDVPDYA 13

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 473505 seqs, 146272329 residues

Total number of hits satisfying chosen parameters: 473505

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_17:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	56	74.7	329	12	O40850	O40850 influenza a
2	56	74.7	329	12	O40764	O40764 influenza a
3	56	74.7	329	12	O40820	O40820 influenza a
4	56	74.7	344	12	Q98083	Q98083 influenza a
5	56	74.7	347	12	Q82787	Q82787 influenza a
6	56	74.7	347	12	Q82790	Q82790 influenza a
7	56	74.7	565	12	Q9WA90	Q9wa90 influenza a
8	55	73.3	236	12	Q9W8Z1	Q9w8z1 influenza a
9	55	73.3	236	12	Q9W8T9	Q9w8t9 influenza a
10	55	73.3	236	12	Q9W8J7	Q9w8j7 influenza a
11	55	73.3	286	12	Q9WI26	Q9wi26 influenza a
12	55	73.3	286	12	Q9WI25	Q9wi25 influenza a
13	55	73.3	286	12	Q9WI24	Q9wi24 influenza a
14	55	73.3	325	12	O40774	O40774 influenza a
15	55	73.3	325	12	Q9DL45	Q9dl45 influenza a
16	55	73.3	326	12	O40779	O40779 influenza a
17	55	73.3	326	12	O40780	O40780 influenza a
18	55	73.3	327	12	O92324	O92324 influenza a
19	55	73.3	327	12	O92325	O92325 influenza a
20	55	73.3	327	12	O92327	O92327 influenza a
21	55	73.3	327	12	Q9YT90	Q9yt90 influenza a
22	55	73.3	327	12	Q9YT89	Q9yt89 influenza a
23	55	73.3	327	12	Q9YT88	Q9yt88 influenza a
24	55	73.3	327	12	Q9YT87	Q9yt87 influenza a
25	55	73.3	327	12	Q9YT86	Q9yt86 influenza a
26	55	73.3	327	12	Q9YRU9	Q9yru9 influenza a
27	55	73.3	327	12	Q9YRU8	Q9yru8 influenza a
28	55	73.3	327	12	Q9YRU7	Q9yru7 influenza a
29	55	73.3	327	12	Q9YRU6	Q9yru6 influenza a
30	55	73.3	327	12	Q9YRU5	Q9yru5 influenza a
31	55	73.3	327	12	Q9YRU4	Q9yru4 influenza a
32	55	73.3	327	12	Q9YRU3	Q9yru3 influenza a
33	55	73.3	327	12	Q9YJ57	Q9yj57 influenza a
34	55	73.3	327	12	Q9YIN9	Q9yin9 influenza a
35	55	73.3	327	12	Q9Q884	Q9q884 influenza a
36	55	73.3	328	12	O40855	O40855 influenza a
37	55	73.3	328	12	O40867	O40867 influenza a
38	55	73.3	328	12	Q82567	Q82567 influenza a
39	55	73.3	328	12	Q82568	Q82568 influenza a
40	55	73.3	328	12	Q82569	Q82569 influenza a
41	55	73.3	328	12	Q82578	Q82578 influenza a
42	55	73.3	328	12	Q82579	Q82579 influenza a
43	55	73.3	328	12	Q82580	Q82580 influenza a
44	55	73.3	328	12	Q82581	Q82581 influenza a
45	55	73.3	328	12	Q82575	Q82575 influenza a

ALIGNMENTS

RESULT 1
O40850
ID O40850 PRELIMINARY; PRT; 329 AA.
AC O40850;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE HEMAGGLUTININ (FRAGMENT).
GN HA.
OS Influenza A virus.
OC Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;
OC Influenza virus A and B group; Influenza A viruses.
OX NCBI_TaxID=11320;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A/WUZHOU/1/94(H3N2);
RA Fitch W.M., Bush R.M., Bender C.A., Cox N.J.;
RL Proc. Natl. Acad. Sci. U.S.A. 0:0-0(1997).
DR EMBL; AF008858; AAB69800.1; -.
DR InterPro; IPR001364; Hemagglutn.
DR Pfam; PF00509; Hemagglutinin; 1.
DR ProDom; PD000225; Hemagglutn; 1.
FT NON_TER 1 1
FT NON_TER 329 329
SQ SEQUENCE 329 AA; 36562 MW; 9DD1CA9187D537DF CRC64;

Query Match 74.7%; Score 56; DB 12; Length 329;
Best Local Similarity 90.9%; Pred. No. 0.29;
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy       3 SGYPYDVPDYA 13
           | |||||||||
Db      96 SCYPYDVPDYA 106
```

RESULT 2
O40764
ID O40764 PRELIMINARY; PRT; 329 AA.
AC O40764;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)

```
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   HEMAGGLUTININ GENE (FRAGMENT).
GN   HA.
OS   Influenza A virus.
OC   Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;
OC   Influenza virus A and B group; Influenza A viruses.
OX   NCBI_TaxID=11320;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=A/HONG_KONG/1/94(H3N2);
RA   Fitch W.M., Bush R.M., Bender C.A., Cox N.J.;
RL   Submitted (JUN-1997) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF008772; AAB63709.1; -.
DR   InterPro; IPR001364; Hemagglutn.
DR   Pfam; PF00509; Hemagglutinin; 1.
DR   ProDom; PD000225; Hemagglutn; 1.
FT   NON_TER       1      1
FT   NON_TER     329    329
SQ   SEQUENCE   329 AA;  36630 MW;  D1DF31668818FA8D CRC64;


  Query Match             74.7%;  Score 56;  DB 12;  Length 329;
  Best Local Similarity   90.9%;  Pred. No. 0.29;
  Matches   10;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      3 SGYPYDVPDYA 13
          | |||||||||
Db     96 SCYPYDVPDYA 106


RESULT   3
O40820
ID   O40820      PRELIMINARY;      PRT;   329 AA.
AC   O40820;
DT   01-JAN-1998 (TrEMBLrel. 05, Created)
DT   01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   HEMAGGLUTININ GENE (FRAGMENT).
GN   HA.
OS   Influenza A virus.
OC   Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;
OC   Influenza virus A and B group; Influenza A viruses.
OX   NCBI_TaxID=11320;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=A/GUANGDONG/25/93(H3N2);
RA   Fitch W.M., Bush R.M., Bender C.A., Cox N.J.;
RL   Submitted (JUN-1997) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF008828; AAB63765.1; -.
DR   InterPro; IPR001364; Hemagglutn.
DR   Pfam; PF00509; Hemagglutinin; 1.
DR   ProDom; PD000225; Hemagglutn; 1.
FT   NON_TER       1      1
FT   NON_TER     329    329
SQ   SEQUENCE   329 AA;  36566 MW;  4DD909217213F423 CRC64;


  Query Match             74.7%;  Score 56;  DB 12;  Length 329;
  Best Local Similarity   90.9%;  Pred. No. 0.29;
  Matches   10;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      3 SGYPYDVPDYA 13
          | |||||||||
Db     96 SCYPYDVPDYA 106


RESULT   4
Q98083
ID   Q98083      PRELIMINARY;      PRT;   344 AA.
AC   Q98083;
DT   01-FEB-1997 (TrEMBLrel. 02, Created)
DT   01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   HAEMAGGLUTININ (FRAGMENT).
OS   Influenza A virus.
OC   Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;
OC   Influenza virus A and B group; Influenza A viruses.
OX   NCBI_TaxID=11320;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=A/GUANDONG/28/94 (H3N2);
RX   MEDLINE=96370797; PubMed=8774693;
RA   Lindstrom S., Sugita S., Endo A., Ishida M., Huang P., Xi S.H.,
RA   Nerome K.;
RT   "Evolutionary characterization of recent human H3N2 influenza A
RT   isolates from Japan and China: novel changes in the receptor binding
RT   domain.";
RL   Arch. Virol. 141:1349-1355(1996).
DR   EMBL; U48442; AAB09416.1; -.
DR   InterPro; IPR001364; Hemagglutn.
DR   Pfam; PF00509; Hemagglutinin; 1.
DR   ProDom; PD000225; Hemagglutn; 1.
FT   CHAIN         17   >344       HAEMAGGLUTININ.
FT   NON_TER      344    344
SQ   SEQUENCE   344 AA;  38139 MW;  C5A8CC7DBB0ACBAE CRC64;


  Query Match             74.7%;  Score 56;  DB 12;  Length 344;
  Best Local Similarity   90.9%;  Pred. No. 0.31;
  Matches   10;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      3 SGYPYDVPDYA 13
          | |||||||||
Db    112 SCYPYDVPDYA 122


RESULT   5
Q82787
ID   Q82787      PRELIMINARY;      PRT;   347 AA.
AC   Q82787;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   HAEMAGGLUTININ (FRAGMENT).
OS   Influenza A virus.
OC   Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;
OC   Influenza virus A and B group; Influenza A viruses.
OX   NCBI_TaxID=11320;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=A/GUANGDONG/25/93;
RX   MEDLINE=96105347; PubMed=7503689;
RA   Ellis J.S., Chakraverty P., Clewley J.P.;
RT   "Genetic and antigenic variation in the haemagglutinin of recently
RT   circulating human influenza A (H3N2) viruses in the United Kingdom.";
RL   Arch. Virol. 140:1889-1904(1995).
DR   EMBL; Z46406; CAA86541.1; -.
DR   InterPro; IPR001364; Hemagglutn.
DR   Pfam; PF00509; Hemagglutinin; 1.
DR   ProDom; PD000225; Hemagglutn; 1.
FT   NON_TER       1      1
FT   NON_TER     347    347
SQ   SEQUENCE   347 AA;  38368 MW;  441AF1B0BE1F594B CRC64;


  Query Match             74.7%;  Score 56;  DB 12;  Length 347;
  Best Local Similarity   90.9%;  Pred. No. 0.31;
  Matches   10;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      3 SGYPYDVPDYA 13
          | |||||||||
Db     96 SCYPYDVPDYA 106


RESULT   6
Q82790
```

```
ID   Q82790      PRELIMINARY;      PRT;   347 AA.
AC   Q82790;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   HAEMAGGLUTININ (FRAGMENT).
OS   Influenza A virus.
OC   Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;
OC   Influenza virus A and B group; Influenza A viruses.
OX   NCBI_TaxID=11320;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=A/HONG KONG/2/94;
RX   MEDLINE=96105347; PubMed=7503689;
RA   Ellis J.S., Chakraverty P., Clewley J.P.;
RT   "Genetic and antigenic variation in the haemagglutinin of recently
RT   circulating human influenza A (H3N2) viruses in the United Kingdom.";
RL   Arch. Virol. 140:1889-1904(1995).
DR   EMBL; Z46408; CAA86543.1; -.
DR   InterPro; IPR001364; Hemagglutn.
DR   Pfam; PF00509; Hemagglutinin; 1.
DR   ProDom; PD000225; Hemagglutn; 1.
FT   NON_TER       1      1
FT   NON_TER     347    347
SQ   SEQUENCE   347 AA;  38379 MW;  C29A4400E80F8635 CRC64;


  Query Match             74.7%;  Score 56;  DB 12;  Length 347;
  Best Local Similarity   90.9%;  Pred. No. 0.31;
  Matches   10;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      3 SGYPYDVPDYA 13
          | |||||||||
Db     96 SCYPYDVPDYA 106


RESULT   7
Q9WA90
ID   Q9WA90      PRELIMINARY;      PRT;   565 AA.
AC   Q9WA90;
DT   01-NOV-1999 (TrEMBLrel. 12, Created)
DT   01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   INFLUENZA A/EQUINE/KENTUCKY/2/86 (H3N8) HEMAGGLUTININ (HA) RNA (SEG.
DE   4), COMPLETE CDS.
OS   Influenza A virus (strain A/Equine/Kentucky/2/86).
OC   Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;
OC   Influenza virus A and B group; Influenza A viruses; Influenza A virus.
OX   NCBI_TaxID=11403;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=89204899; PubMed=2705299;
RA   Kawaoka Y., Bean W.J., Webster R.G.;
RT   "Evolution of the hemagglutinin of equine H3 influenza viruses.";
RL   Virology 169:283-292(1989).
DR   EMBL; M24727; AAA43102.1; -.
DR   InterPro; IPR001364; Hemagglutn.
DR   Pfam; PF00509; Hemagglutinin; 1.
DR   PRINTS; PR00329; HEMAGGLUTN12.
DR   ProDom; PD000225; Hemagglutn; 1.
FT   CHAIN         1    344       POTENTIAL.
FT   CHAIN       345    565       POTENTIAL.
SQ   SEQUENCE   565 AA;  63592 MW;  EB9FEFF7CBB861DB CRC64;


  Query Match             74.7%;  Score 56;  DB 12;  Length 565;
  Best Local Similarity   69.2%;  Pred. No. 0.55;
  Matches    9;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 SGSGYPYDVPDYA 13
          | : ||||:||||
Db    109 SSNCYPYDIPDYA 121


RESULT   8
Q9W8Z1
ID   Q9W8Z1      PRELIMINARY;      PRT;   236 AA.
AC   Q9W8Z1;
DT   01-NOV-1999 (TrEMBLrel. 12, Created)
DT   01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   HEMAGGLUTININ (FRAGMENT).
OS   Influenza A virus.
OC   Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;
OC   Influenza virus A and B group; Influenza A viruses.
OX   NCBI_TaxID=11320;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=TOKYO1527, TOKYO1511;
RA   Mori S., Sekine H.;
RT   "HA1 domain of Influenza A (H3N2) virus.";
RL   Submitted (MAY-1998) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AB013807; BAA77285.1; -.
DR   EMBL; AB013806; BAA77284.1; -.
DR   InterPro; IPR001364; Hemagglutn.
DR   Pfam; PF00509; Hemagglutinin; 1.
DR   PRINTS; PR00329; HEMAGGLUTN12.
DR   ProDom; PD000225; Hemagglutn; 1.
FT   NON_TER       1      1
FT   NON_TER     236    236
SQ   SEQUENCE   236 AA;  26477 MW;  4566C8E7210FB558 CRC64;


  Query Match             73.3%;  Score 55;  DB 12;  Length 236;
  Best Local Similarity   100.0%;  Pred. No. 0.29;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      5 YPYDVPDYA 13
          |||||||||
Db     11 YPYDVPDYA 19


RESULT   9
Q9W8T9
ID   Q9W8T9      PRELIMINARY;      PRT;   236 AA.
AC   Q9W8T9;
DT   01-NOV-1999 (TrEMBLrel. 12, Created)
DT   01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   HEMAGGLUTININ (FRAGMENT).
OS   Influenza A virus.
OC   Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;
OC   Influenza virus A and B group; Influenza A viruses.
OX   NCBI_TaxID=11320;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=TOKYO1570, TOKYO1566, TOKYO1567, TOKYO1568, TOKYO1569;
RA   Mori S., Sekine H.;
RT   "HA1 domain of Influenza A (H3N2) virus.";
RL   Submitted (MAY-1998) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AB013813; BAA77291.1; -.
DR   EMBL; AB013809; BAA77287.1; -.
DR   EMBL; AB013810; BAA77288.1; -.
DR   EMBL; AB013811; BAA77289.1; -.
DR   EMBL; AB013812; BAA77290.1; -.
DR   InterPro; IPR001364; Hemagglutn.
DR   Pfam; PF00509; Hemagglutinin; 1.
DR   PRINTS; PR00329; HEMAGGLUTN12.
DR   ProDom; PD000225; Hemagglutn; 1.
FT   NON_TER       1      1
FT   NON_TER     236    236
SQ   SEQUENCE   236 AA;  26491 MW;  44CDD257210FB558 CRC64;


  Query Match             73.3%;  Score 55;  DB 12;  Length 236;
  Best Local Similarity   100.0%;  Pred. No. 0.29;
```

```
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       5 YPYDVPDYA 13
           |||||||||
Db      11 YPYDVPDYA 19


RESULT  10
Q9W8J7
ID   Q9W8J7         PRELIMINARY;   PRT;   236 AA.
AC   Q9W8J7;
DT   01-NOV-1999 (TrEMBLrel. 12, Created)
DT   01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   HEMAGGLUTININ (FRAGMENT).
GN   HA.
OS   Influenza A virus.
OC   Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;
OC   Influenza virus A and B group; Influenza A viruses.
OX   NCBI_TaxID=11320;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=72, 70, AND 71;
RA   Nagashima M., Mori S., Sekine H.;
RT   "Influenza A virus gene for hemagglutinin.";
RL   Submitted (MAY-1998) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=TOKYO1539;
RA   Mori S., Sekine H.;
RT   "HA1 domain of Influenza A (H3N2) virus.";
RL   Submitted (MAY-1998) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AB014062; BAA77294.1; -.
DR   EMBL; AB013808; BAA77286.1; -.
DR   EMBL; AB014060; BAA77292.1; -.
DR   EMBL; AB014061; BAA77293.1; -.
DR   InterPro; IPR001364; Hemagglutn.
DR   Pfam; PF00509; Hemagglutinin; 1.
DR   PRINTS; PR00329; HEMAGGLUTN12.
DR   ProDom; PD000225; Hemagglutn; 1.
FT   NON_TER       1      1
FT   NON_TER     236    236
SQ   SEQUENCE   236 AA;  26493 MW;  457455682C8D28D7 CRC64;


  Query Match             73.3%;  Score 55;  DB 12;  Length 236;
  Best Local Similarity   100.0%;  Pred. No. 0.29;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       5 YPYDVPDYA 13
           |||||||||
Db      11 YPYDVPDYA 19


RESULT  11
Q9WI26
ID   Q9WI26      PRELIMINARY;      PRT;   286 AA.
AC   Q9WI26;
DT   01-NOV-1999 (TrEMBLrel. 12, Created)
DT   01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   H3 HEMAGGLUTININ (FRAGMENT).
GN   HA.
OS   influenza A virus (A/Taiwan/2034/96(H3N2)).
OC   Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;
OC   Influenza virus A and B group; Influenza A viruses; Influenza A virus.
OX   NCBI_TaxID=95241;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=A/TAIWAN/2034/96;
RA   Shih S.-R., Chen S.-H., Wu S.-L., Ho C.-C.;
RT   "Sequence Analysis of HA and NA Genes of Influenza Viruses Isolated in
RT   Taiwan.";
RL   Submitted (APR-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF139937; AAD34854.1; -.
DR   InterPro; IPR001364; Hemagglutn.
DR   Pfam; PF00509; Hemagglutinin; 1.
DR   PRINTS; PR00329; HEMAGGLUTN12.
DR   ProDom; PD000225; Hemagglutn; 1.
FT   NON_TER       1      1
FT   NON_TER     286    286
SQ   SEQUENCE   286 AA;  32075 MW;  7BA39C7632D33186 CRC64;


  Query Match             73.3%;  Score 55;  DB 12;  Length 286;
  Best Local Similarity   100.0%;  Pred. No. 0.36;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       5 YPYDVPDYA 13
           |||||||||
Db      57 YPYDVPDYA 65


RESULT  12
Q9WI25
ID   Q9WI25      PRELIMINARY;      PRT;   286 AA.
AC   Q9WI25;
DT   01-NOV-1999 (TrEMBLrel. 12, Created)
DT   01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   H3 HEMAGGLUTININ (FRAGMENT).
GN   HA.
OS   influenza A virus (A/Taiwan/2192/96(H3N2)).
OC   Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;
OC   Influenza virus A and B group; Influenza A viruses; Influenza A virus.
OX   NCBI_TaxID=95240;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=A/TAIWAN/2192/96;
RA   Shih S.-R., Chen S.-H., Wu S.-L., Ho C.-C.;
RT   "Sequence Analysis of HA and NA Genes of Influenza Viruses Isolated in
RT   Taiwan.";
RL   Submitted (APR-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF139936; AAD34853.1; -.
DR   InterPro; IPR001364; Hemagglutn.
DR   Pfam; PF00509; Hemagglutinin; 1.
DR   PRINTS; PR00329; HEMAGGLUTN12.
DR   ProDom; PD000225; Hemagglutn; 1.
FT   NON_TER       1      1
FT   NON_TER     286    286
SQ   SEQUENCE   286 AA;  32059 MW;  CB51B092AAD0D7D8 CRC64;


  Query Match             73.3%;  Score 55;  DB 12;  Length 286;
  Best Local Similarity   100.0%;  Pred. No. 0.36;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       5 YPYDVPDYA 13
           |||||||||
Db      57 YPYDVPDYA 65


RESULT  13
Q9WI24
ID   Q9WI24      PRELIMINARY;      PRT;   286 AA.
AC   Q9WI24;
DT   01-NOV-1999 (TrEMBLrel. 12, Created)
DT   01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   H3 HEMAGGLUTININ (FRAGMENT).
GN   HA.
OS   influenza A virus (A/Taiwan/2191/96(H3N2)).
OC   Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;
OC   Influenza virus A and B group; Influenza A viruses; Influenza A virus.
OX   NCBI_TaxID=95236;
RN   [1]
```

```
RP   SEQUENCE FROM N.A.
RC   STRAIN=A/TAIWAN/2191/96;
RA   Shih S.-R., Chen S.-H., Wu S.-L., Ho C.-C.;
RT   "Sequence Analysis of HA and NA Genes of Influenza Viruses Isolated in
RT   Taiwan.";
RL   Submitted (APR-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF139932; AAD34849.1; -.
DR   InterPro; IPR001364; Hemagglutn.
DR   Pfam; PF00509; Hemagglutinin; 1.
DR   PRINTS; PR00329; HEMAGGLUTN12.
DR   ProDom; PD000225; Hemagglutn; 1.
FT   NON_TER       1      1
FT   NON_TER     286    286
SQ   SEQUENCE   286 AA;  31976 MW;  69AB2AE727D26078 CRC64;


  Query Match             73.3%;  Score 55;  DB 12;  Length 286;
  Best Local Similarity   100.0%;  Pred. No. 0.36;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       5 YPYDVPDYA 13
           |||||||||
Db      57 YPYDVPDYA 65


RESULT  14
O40774
ID   O40774      PRELIMINARY;      PRT;   325 AA.
AC   O40774;
DT   01-JAN-1998 (TrEMBLrel. 05, Created)
DT   01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   HEMAGGLUTININ GENE (FRAGMENT).
GN   HA.
OS   Influenza A virus.
OC   Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;
OC   Influenza virus A and B group; Influenza A viruses.
OX   NCBI_TaxID=11320;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=A/ILLINOIS/5/95(H3N2);
RA   Fitch W.M., Bush R.M., Bender C.A., Cox N.J.;
RL   Submitted (JUN-1997) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF008782; AAB63719.1; -.
DR   InterPro; IPR001364; Hemagglutn.
DR   Pfam; PF00509; Hemagglutinin; 1.
DR   ProDom; PD000225; Hemagglutn; 1.
FT   NON_TER       1      1
FT   NON_TER     325    325
SQ   SEQUENCE   325 AA;  36103 MW;  4E5B4648C2024765 CRC64;


  Query Match             73.3%;  Score 55;  DB 12;  Length 325;
  Best Local Similarity   100.0%;  Pred. No. 0.42;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       5 YPYDVPDYA 13
           |||||||||
Db      98 YPYDVPDYA 106


RESULT  15
Q9DL45
ID   Q9DL45      PRELIMINARY;      PRT;   325 AA.
AC   Q9DL45;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   HEMAGGLUTININ (FRAGMENT).
OS   Influenza A virus (A/Finland/583/98(H3N2)).
OC   Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;
OC   Influenza virus A and B group; Influenza A viruses; Influenza A virus.
OX   NCBI_TaxID=147149;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=A/FINLAND/583/98;
RA   Pyhala R., Haanpaa M., Kleemola M., Tervahauta R., Visakorpi R.,
RA   Kinnunen L.;
RT   "Acceptable protective efficacy of influenza vaccination in young
RT   military conscripts in circumstances of incomplete antigenic and
RT   genetic match.";
RL   Vaccine 0:0-0(2001).
DR   EMBL; AF311689; AAG47810.1; -.
DR   InterPro; IPR001364; Hemagglutn.
DR   Pfam; PF00509; Hemagglutinin; 1.
DR   PRINTS; PR00329; HEMAGGLUTN12.
DR   ProDom; PD000225; Hemagglutn; 1.
FT   NON_TER       1      1
FT   NON_TER     325    325
SQ   SEQUENCE   325 AA;  36029 MW;  5BECB8CC3E83B1C2 CRC64;


  Query Match             73.3%;  Score 55;  DB 12;  Length 325;
  Best Local Similarity   100.0%;  Pred. No. 0.42;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       5 YPYDVPDYA 13
           |||||||||
Db      98 YPYDVPDYA 106


Search completed: April 11, 2002, 10:22:57
Job time: 427 sec
```

THIS PAGE BLANK (USPTO)